

UTILITY PATENT APPLICATION TRANSMITTAL (Only for new nonprovisional applications under 37 CFR 1 53(b)) Title

Attorney Docket No.	4085-226-27

First Inventor or Application Identifier

Michelle A.J. PALMER, et al

RECEPTOR FUNCTION ASSAY FOR G-PROTEIN COUPLED RECEPTORS AND ORPHAN RECEPTORS BY REPORTER ENZYME MUTANT COMPLEMENTATION

APPLICATION ELEMENTS See MPEP chapter 600 concerning utility patent application contents	Assistant Commissioner for Patents ADDRESS TO: Box Patent Application Washington, DC 20231					
Fee Transmittal Form (e.g. PTO/SB/17) (Submit an onginal and a duplicate for fee processing)	ACCOMPANYING APPLICATION PARTS					
2. Specification Total Pages 25	 6. □ Assignment Papers (cover sheet & document(s)) 7. □ 37 C.F.R. §3.73(b) Statement □ Power of Attorney (when there is an assignee) 8. □ English Translation Document (if applicable) 					
3.	9. Information Disclosure Statement (IDS)/PTO-1449 Dreliminary Amendment					
4. □ Oath or Declaration Total Pages	11. White Advance Serial No. Postcard					
a. □ Newly executed (original or copy) b. □ Copy from a prior application (37 C.F.R. §1.63(d)) (for continuation/divisional with box 15 completed)	12. Small Entity Statement(s) Statement filed in prior application. Status still proper and desired.					
i. □ DELETION OF INVENTOR(S) Signed statement attached deleting inventor(s) named	13. Certified Copy of Priority Document(s) (if foreign priority is claimed)					
in the prior application, see 37 C.F.R. §1.63(d)(2) and 1.33(b).	14. ■ Other: List of Inventors' Names and Addresses					
5. Incorporation By Reference (usable if box 48 is checked) The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4B, is considered to be part of the disclosure of the accompanying application and is hereby incorporated by reference therein.						
15. If a CONTINUING APPLICATION, check appropriate box, and supp						
☐ Continuation ☐ Divisional ☐ Continuation-i Prior application information: Examiner:	n-part (CIP) of prior application no.: Group Art Unit:					
16. Amend the specification by inserting before the first line the se	·					
□ This application is a □ Continuation □ Division □ Continuation-in-part (CIP) of application Serial No. Filed on						
■ This application claims priority of provisional application Serial No. 60/180,669 Filed February 7, 2000						
17. CORRESPONDENCE ADDRESS						
Steven B. Kelber PIPER MARBURY RUDNICK & WOLFE LLP 1200 Nineteenth Street, N.W. Washington, D.C. 20036-2412 Telephone No. (202) 861-3900 Facsimile No. (202) 223-2085						

Name:	Steven B. Kelber	Registrati	on No.:	30,073
Signature:	Be we		Date:	9/1/00
Name:	Laura A. Donnelly	Registrati	on No.:	38,435

Docket No.

4085-226-27

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

INVENTOR(S) Michelle A.J. PALMER, et al

SERIAL NO:

New Application

FILING DATE: Herewith

FOR:

ij 13

The first start start start of the start s

RECEPTOR FUNCTION ASSAY FOR G-PROTEIN COUPLED RECEPTORS AND ORPHAN

RECEPTORS BY REPORTER ENZYME MUTANT COMPLEMENTATION

FEE TRANSMITTAL

ASSISTANT COMMISSIONER FOR PATENTS WASHINGTON, D.C. 20231

FOR	NUMBER FILED	NUMBER EXTRA	RATE	CALCULATIONS
TOTAL CLAIMS	30 - 20 =	10	× \$18 =	\$180.00
INDEPENDENT CLAIMS	8 - 3 =	5	× \$78 =	\$390.00
□ MULTIPLE DEPENDEN	\$0.00			
■ LATE FILING OF DECI	\$130.00			
	\$690.00			
	\$1,390.00			
☐ REDUCTION BY 50% F	\$0.00			
□ FILING IN NON-ENGLISH LANGUAGE			+ \$130 =	\$0.00
RECORDATION OF ASSIGNMENT			+ \$40 =	\$0.00
			TOTAL	\$1,390.00

Please charge Deposit Account No. 50-1442 in the amount of

. A duplicate copy of this sheet is enclosed.

A check in the amount of

\$1,390.00

to cover the filing fee is enclosed.

The Commissioner is hereby authorized to charge any additional fees which may be required for the papers being filed herewith and for which no check is enclosed herewith, or credit any overpayment to Deposit Account No. 50-1442. A duplicate copy of this sheet is enclosed.

Date:

1200 Nineteenth Street, N.W. Washington, D.C. 20036-2412 Telephone No. (202) 861-3900 Facsimile No. (202) 223-2085

Respectfully submitted,

PIPER MARBURY RUDNICK & WOLFE LLP

Steren B. Kelber

Registration No. 30,073

Laura A. Donnelly

Registration No. 38,435

TITLE OF THE INVENTION

RECEPTOR FUNCTION ASSAY FOR G-PROTEIN COUPLED RECEPTORS AND ORPHAN RECEPTORS BY REPORTER ENZYME MUTANT COMPLEMENTATION

BACKGROUND OF THE INVENTION

This application claims the benefit from Provisional Application Serial No. 60/180,669, filed February 7, 2000. The entirety of that provisional application is incorporated herein by reference.

Field of the Invention

This invention relates to methods of detecting G-protein-coupled receptor (GPCR) activity, and provides methods of assaying GPCR activity and methods for screening for GPCR ligands, G-protein-coupled receptor kinase (GRK) activity, and compounds that interact with components of the GPCR regulatory process.

The actions of many extracellular signals are mediated by the interaction of G-protein-coupled receptors (GPCRs) and guanine nucleotide-binding regulatory proteins (G-proteins). G-protein-mediated signaling systems have been identified in many divergent organisms, such as mammals and yeast. The GPCRs represent a large super family of proteins which have divergent amino acid sequences, but share common structural features, in particular, the presence of seven transmembrane helical domains. GPCRs respond to, among other extracellular signals, neurotransmitters, hormones, odorants and light. Individual GPCR types activate a particular signal transduction pathway; at least ten different signal transduction pathways are known to be activated via GPCRs. For example, the beta 2-adrenergic receptor (β 2AR) is a prototype mammalian GPCR. In response to agonist binding, β 2AR receptors activate a G-protein (Gs) which in turn stimulates adenylate cyclase activity and results in increased cyclic adenosine monophosphate (cAMP) production in the cell.

20

25

5

5

The signaling pathway and final cellular response that result from GPCR stimulation depends on the specific class of G-protein with which the particular receptor is coupled (Hamm, "The many faces of G-Protein Signaling." J. Biol. Chem., 273:669-672 (1998)). For instance, coupling to the Gs class of G-proteins stimulates cAMP production and activation of Protein Kinase A and C pathways, whereas coupling to the Gi class of G-proteins down regulates cAMP. Other second messenger systems as calcium, phosphlipase C, and phosphatidylinositol 3 may also be utilized. As a consequence, GPCR signaling events have predominantly been measured via quantification of these second messenger products.

A common feature of GPCR physiology is desensitization and recycling of the receptor through the processes of receptor phosphorylation, endocytosis and dephosphorylation (Ferguson, et al., "G-protein-coupled receptor regulation: role of G-protein-coupled receptor kinases and arrestins." Can. J. Physiol. Pharmacol., 74:1095-1110 (1996)). Ligand-occupied GPCRs can be phosphorylated by two families of serine/threonine kinases, the G-protein-coupled receptor kinases (GRKs) and the second messenger-dependent protein kinases such as protein kinase A and protein kinase C. Phosphorylation by either class of kinases serves to down-regulate the receptor by uncoupling it from its corresponding G-protein. GRK-phosphorylation also serves to down-regulate the receptor by recruitment of a class of proteins known as the arrestins that bind the cytoplasmic domain of the receptor and promote clustering of the receptor into endocytic vescicles. Once the receptor is endocytosed, it will either be degraded in lysosomes or dephosphorylated and recycled back to the plasma membrane as fully-functional receptor.

Binding of an arrestin protein to an activated receptor has been documented as a common phenomenon for a variety of GPCRs ranging from rhodopsin to β2AR to the

5

neurotensin receptor (<u>Barak</u>, et al., "A β-arrestin/Green Fluorescent fusion protein biosensor for detecting G-Protein-Coupled Receptor Activation," J. Biol. Chem., 272:27497-500 (1997)). Consequently, monitoring arrestin interaction with a specific GPCR can be utilized as a generic tool for measuring GPCR activation. Similarly, a single G-protein and GRK also partner with a variety of receptors (<u>Hamm</u>, et al. (1998) and <u>Pitcher et al.</u>, "G-Protein-Coupled Receptor Kinases," Annu. Rev. Biochem., 67:653-92 (1998)), such that these protein/protein interactions may also be monitored to determine receptor activity.

The present invention involves the use of a proprietary technology (ICASTTM,
Intercistronic Complementation Analysis Screening Technology) for monitoring
protein/protein interactions in GPCR signaling. The method involves using two inactive βgalactosidase mutants, each of which is fused with one of two interacting protein pairs, such
as a GPCR and an arrestin. The formation of an active β-galactosidase complex is driven by
interaction of the target proteins. In this system, β-galactosidase activity acts as a read out of
GPCR activity. FIGURE 23 is a schematic depicting the method of the present invention.
FIGURE 23 shows two inactive mutants that become active when they interact. In addition,
this technology could be used to monitor GPCR-mediated signaling pathways via other
downstream signaling components such as G-proteins, GRKs or c-Src.

Many therapeutic drugs in use today target GPCRs, as they regulate vital physiological responses, including vasodilation, heart rate, bronchodilation, endocrine secretion and gut peristalsis. See, e.g., Lefkowitz et al., Annu. Rev. Biochem., 52:159 (1983). For instance, drugs targeting the highly studied GPCR, β2AR are used in the treatment of anaphylaxis, shock hypertension, asthma and other conditions. Some of these drugs mimic

20

5

the ligand for this receptor. Other drugs act to antagonize the receptor in cases when disease arises from spontaneous activity of the receptor.

Efforts such as the Human Genome Project are identifying new GPCRs ("orphan" receptors) whose physiological roles and ligands are unknown. It is estimated that several thousand GPCRs exist in the human genome. Of the 250 GPCRs identified to date, only 150 have been associated with ligands.

SUMMARY OF THE INVENTION

A first aspect of the present invention is a method that monitors GPCR function proximally at the site of receptor activation, thus providing more information for drug discovery purposes due to fewer competing mechanisms. Activation of the GPCR is measured by a read-out for interaction of the receptor with a regulatory component such as arrestin, G-protein, GRK or other kinases, the binding of which to the receptor is dependent upon agonist occupation of the receptor. Protein/protein interaction is detected by complementation of reporter proteins such as utilized by the ICASTTM technology.

A further aspect of the present invention is a method of assessing G-protein-coupled receptor (GPCR) pathway activity under test conditions by providing a test cell that expresses a GPCR, e.g., muscarinic, adrenergic, dopamine, angiotensin or endothelin, as a fusion protein to a mutant reporter protein and interacting, i.e., G-proteins, arrestin or GRK, as a fusion protein with a complementing reporter protein. When test cells are exposed to a known agonist to the target GPCR under test conditions, activation of the GPCR will be monitored by complementation of the reporter enzyme. Increased reporter enzyme activity reflects interaction of the GPCR with its interacting protein partner.

5

A further aspect of the present invention is a method of assessing GPCR pathway activity in the presence of a test kinase.

A further aspect of the present invention is a method of assessing GPCR pathway activity in the presence of a test G-protein.

A further aspect of the present invention is a method of assessing GPCR pathway activity upon exposure of the test cell to a test ligand.

A further aspect of the present invention is a method of assessing GPCR pathway activity upon co-expression in the test cell of a second receptor.

A further aspect of the present invention is a method for screening for a ligand or agonists to an orphan GPCR. The ligand or agonist could be contained in natural or synthetic libraries or mixtures or could be a physical stimulus. A test cell is provided that expresses the orphan GPCR as a fusion protein with one β -galactosidase mutant and, for example, an arrestin or mutant form of arrestin as a fusion protein with another β -galactosidase mutant. The interaction of the arrestin with the orphan GPCR upon receptor activation is measured by enzymatic activity of the complemented β -galactosidase. The test cell is exposed to a test compound, and an increase in β -galactosidase activity indicates the presence of a ligand or agonist.

A further aspect of the present invention is a method for screening a protein of interest, for example, an arrestin protein (or mutant form of the arrestin protein) for the ability to bind to a phosphorylated, or activated, GPCR. A cell is provided that expresses a GPCR and contains β -arrestin. The cell is exposed to a known GPCR agonist and then reporter enzyme activity is detected. Increased reporter enzyme activity indicates that the β -arrestin molecule can bind to phosphorylated, or activated, GPCR in the test cell.

20

5

A further aspect of the present invention is a method to screen for an agonist to a specific GPCR. The agonist could be contained in natural or synthetic libraries or could be a physical stimulus. A test cell is provided that expresses a GPCR as a fusion protein with one β -galactosidase mutant and, for example, an arrestin as a fusion protein with another β -galactosidase mutant. The interaction of arrestin with the GPCR upon receptor activation is measured by enzymatic activity of the complemented β -galactosidase. The test cell is exposed to a test compound, and an increase in β -galactosidase activity indicates the presence of an agonist. The test cell may express a known GPCR or a variety of known GPCRs, or may express an unknown GPCR or a variety of unknown GPCRs. The GPCR may be, for example, an odorant GPCR or a β AR GPCR.

A further aspect of the present invention is a method of screening a test compound for G-protein-coupled receptor (GPCR) antagonist activity. A test cell is provided that expresses a GPCR as a fusion protein with one β -galactosidase mutant and, for example, an arrestin as a fusion protein with another β -galactosidase mutant. The interaction of arrestin with the GPCR upon receptor activation is measured by enzymatic activity of the complemented β -galactosidase. The test cell is exposed to a test compound, and an increase in β -galactosidase activity indicates the presence of an agonist. The cell is exposed to a test compound and to a GPCR agonist, and reporter enzyme activity is detected. When exposure to the agonist occurs at the same time as or subsequent to exposure to the test compound, a decrease in β -galactosidase activity after exposure to the test compound indicates that the test compound has antagonist activity to the GPCR.

A further aspect of the present invention is a method of screening a sample solution for the presence of an agonist, antagonist or ligand to a G-protein-coupled receptor (GPCR).

5

A test cell is provided that expresses a GPCR fusion and contains, for example, a β-arrestin protein fusion. The test cell is exposed to a sample solution, and reporter enzyme activity is assessed. Changed reporter enzyme activity after exposure to the sample solution indicates the sample solution contains an agonist, antagonist or ligand for a GPCR expressed in the cell.

A further aspect of the present invention is a method of screening a cell for the presence of a G-protein-coupled receptor (GPCR).

A further aspect of the present invention is a method of screening a plurality of cells for those cells which contain a G-protein coupled receptor (GPCR).

A further aspect of the invention is a method for mapping GPCR-mediated signaling pathways. For instance, the system could be utilized to monitor interaction of c-src with βarrestin-1 upon GPCR activation. Additionally, the system could be used to monitor protein/protein interactions involved in cross-talk between GPCR signaling pathways and other pathways such as that of the receptor tyrosine kinases or Ras/Raf.

A further aspect of the invention is a method for monitoring homo- or heterodimerization of GPCRs upon agonist or antagonist stimulation.

A further aspect of the invention is a method of screening a cell for the presence of a G-protein-coupled receptor (GPCR) responsive to a GPCR agonist. A cell is provided that contains protein partners that interact downstream in the GPCR's pathway. The protein partners are expressed as fusion proteins to the mutant, complementing enzyme and are used to monitor activation of the GPCR. The cell is exposed to a GPCR agonist and then enzymatic activity of the reporter enzyme is detected. Increased reporter enzyme activity indicates that the cell contains a GPCR responsive to the agonist.

20

The invention is achieved by using ICASTTM protein/protein interaction screening to map signaling pathways. This technology is applicable to a variety of known and unknown GPCRs with diverse functions. They include, but are not limited to, the following subfamilies of GPCRs:

5

- (a) receptors that bind to amine-like ligands-Acetylcholine muscarinic receptor (M1 to M5), alpha and beta Adrenoceptors, Dopamine receptors (D1, D2, D3 and D4), Histamine receptors (H1 and H2), Octopamine receptor and Serotonin receptors (5HT1, 5HT2, 5HT4, 5HT5, 5HT6, 5HT7);
- (b) receptors that bind to a peptide ligand-Angiotensin receptor, Bombesin receptor, Bradykinin receptor, C-C chemokine receptors (CCR1 to CCR8, and CCR10), C-X-C type Chemokine receptors (CXC-R5), Cholecystokinin type A receptor, CCK type receptors, Endothelin receptor, Neurotesin receptor, FMLP-related receptors, Somatostatin receptors (type 1 to type 5) and Opioid receptors (type D, K, M, X);
- (c) receptors that bind to hormone proteins- Follic stimulating hormone receptor,

 Thyrotrophin receptor and Lutropin-choriogonadotropic hormone receptor;
- (d) receptors that bind to neurotransmitters-substance P receptor, Substance K receptor and neuropeptide Y receptor;
 - (e) Olfactory receptors-Olfactory type 1 to type 11, Gustatory and odorant receptors;
- (f) Prostanoid receptors-Prostaglandin E2 (EP1 to EP4 subtypes), Prostacyclin and Thromboxane;
- (g) receptors that bind to metabotropic substances-Metabotropic glutamate group I to group III receptors;

20

5

- (h) receptors that respond to physical stimuli, such as light, or to chemical stimuli, such as taste and smell; and
 - (i) orphan GPCRs-the natural ligand to the receptor is undefined.

ICASTTM provides many benefits to the screening process, including the ability to monitor protein interactions in any sub-cellular compartment-membrane, cytosol and nucleus; the ability to achieve a more physiologically relevant model without requiring protein overexpression; and the ability to achieve a functional assay for receptor binding allowing high information content.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1. Cellular expression levels of $\beta 2$ adrenergic receptor ($\beta 2AR$) and β -arrestin-2 ($\beta Arr2$) in C2 clones. Quantification of β -gal fusion protein was performed using antibodies against β -gal and purified β -gal protein in a titration curve by a standardized ELISA assay. Figure 1A shows expression levels of $\beta 2AR$ - $\beta gal\Delta\alpha$ clones (in expression vector pICAST ALC). Figure 1B shows expression levels of $\beta Arr2$ - $\beta gal\Delta\alpha$ in expression vector pICAST OMC4 for clones 9-3, -7, -9, -10, -19 and -24, or in expression vector pICAST OMN4 for clones 12-4, -9, -16, -18, -22 and -24.

FIGURE 2. Receptor β2AR activation was measured by agonist-stimulated cAMP production. C2 cells expressing pICAST ALC β2AR (clone 5) or parental cells were treated with increasing concentrations of (-)isoproterenol and 0.1mM IBMX. The quantification of cAMP level was expressed as pmol/well.

20

5

FIGURE 3. Interaction of activated receptor $\beta 2AR$ and arrestin can be measured by β -galactosidase complementation. Figure 3A shows a time course of β -galactosidase activity in response to agonist (-)isoproterenol stimulation in C2 expressing $\beta 2AR$ - $\beta gal\Delta\alpha$ ($\beta 2AR$ alone, in expression vector pICAST ALC), or C2 clones, and a pool of C2 co-expressing $\beta 2AR$ - $\beta gal\Delta\alpha$ and $\beta Arr2$ - $\beta gal\Delta\omega$ (in expression vectors pICAST ALC and pICAST OMC). Figure 3B shows a time course of β galactosidase activity in response to agonist (-)isoproterenol stimulation in C2 cells expressing $\beta 2AR$ alone (in expression vector pICAST ALC) and C2 clones co-expressing $\beta 2AR$ and $\beta Arr1$ (in expression vectors ICAST ALC and pICAST OMC).

FIGURE 4. Agonist dose response for interaction of β 2AR and arrestin can be measured by β -galactosidase complementation. Figure 4A shows a dose response to agonists (-)isoproterenol and procaterol in C2 cells co-expressing pICAST ALC β 2AR and pICAST OMC β Arr2 fusion constructs. Figure 4B shows a dose response to agonists (-)isoproterenol and procaterol in C2 cells co-expressing pICAST ALC β 2AR and pICAST OMC β Arr1 fusion constructs.

FIGURE 5. Antagonist mediated inhibition of receptor activity can be measured by β -galactosidase complementation in cells co-expressing β 2AR- β gal $\Delta\alpha$ and β Arr- β gal $\Delta\omega$. Figure 5A shows specific inhibition with adrenergic antagonists ICI-118,551 and propranolol of β -galactosidase activity in C2 clones co-expressing pICAST ALC β 2AR and pICAST OMC β Arr2 fusion constructs after incubation with agonist (-)isoproterenol. Figure 5B shows specific inhibition of β -galactosidase activity with adrenergic antagonists ICI-118,551

20

5

and propranolol in C2 clones co-expressing pICAST ALC β2AR and pICAST OMC βArr1 fusion constructs in the presence of agonist (-)isoproterenol.

FIGURE 6. C2 cells expressing adenosine receptor A2a show cAMP induction in response to agonist (CGC-21680) treatment. C2 parental cells and C2 cells co-expressing pICAST ALC A2aR and pICAST OMC βArr1 as a pool or as selected clones were measured for agonist-induced cAMP response (pmol/well).

FIGURE 7. Agonist stimulated cAMP response in C2 cells co-expressing Dopamine receptor D1 (D1-βgalΔα) and β-arrestin-2 (βArr2-βgalΔω). The clone expressing βArr2-βgalΔω (Arr2 alone) was used as a negative control in the assay. Cells expressing D1-βgalΔα in addition to βArr2-βgalΔω responded agonist treatment (3-hydroxytyramine hydrochloride at 3 μM). D1(PIC2) or D1(PIC3) designate D1 in expression vector pICAST ALC2 or pICAST ALC4, respectively.

FIGURE 8. Variety of mammalian cell lines can be used to generate stable cells for monitoring GPCR and arrestin interactions. FIGURE 8A, FIGURE 8B and FIGURE 8C show the examples of HEK293, CHO and CHW cell lines co-expressing adrenergic receptor β 2AR and arrestin fusion proteins of β -galactosidase mutants. The β -galactosidase activity was used to monitor agonist-induced interaction of β 2AR and arrestin proteins.

FIGURE 9. Beta-gal complementation can be used to monitor $\beta2$ adrenergic receptor homo-dimerization. FIGURE 9A shows β -galactosidase activity in HEK293 clones coexpressing pICAST ALC $\beta2AR$ and pICAST OMC $\beta2AR$. FIGURE 9B shows a cAMP response to agonist (-)isoproterenol in HEK 293 clones co-expressing pICAST ALC $\beta2AR$

20

5

and pICAST OMC β 2AR. HEK293 parental cells were included in the assays as negative controls.

FIGURE 10A. pICAST ALC: Vector for expression of β -gal $\Delta\alpha$ as a C-terminal fusion to the target protein. This construct contains the following features: MCS, multiple cloning site for cloning the target protein in frame with the β -gal $\Delta\alpha$; GS Linker, (GGGGS)n; NeoR, neomycin resistance gene; IRES, internal ribosome entry site; ColE1ori, origin of replication for growth in E. coli; 5'MoMuLV LTR and 3'MoMuLV LTR, viral promotor and polyadenylation signals from the Moloney Murine leukemia virus.

FIGURE 10B. Nucleotide sequence for pICAST ALC.

FIGURE 11A. pICAST ALN: Vector for expression of β -gal $\Delta\alpha$ as an N-terminal fusion to the target protein. This construct contains the following features: MCS, multiple cloning site for cloning the target protein in frame with the β -gal $\Delta\alpha$; GS Linker, (GGGGS)n; NeoR, neomycin resistance gene; IRES, internal ribosome entry site; ColE1ori, origin of replication for growth in E. coli; 5'MoMuLV LTR and 3'MoMuLV LTR, viral promotor and polyadenylation signals from the Moloney Murine leukemia virus.

FIGURE 11B. Nucleotide sequence for pICAST ALN.

FIGURE 12A. pICAST OMC: Vector for expression of β -gal $\Delta\omega$ as a C-terminal fusion to the target protein. This construct contains the following features: MCS, multiple cloning site for cloning the target protein in frame with the β -gal $\Delta\omega$; GS Linker, (GGGGS)n; Hygro, hygromycin resistance gene; IRES, internal ribosome entry site; ColE1ori, origin of replication for growth in E. coli; 5'MoMuLV LTR and 3'MoMuLV LTR, viral promotor and polyadenylation signals from the Moloney Murine leukemia virus.

20

5

FIGURE 12B. Nucleotide sequence for pICAST OMC.

FIGURE 13A. pICAST OMN: Vector for expression of β-galΔω as an N-terminal fusion to the target protein. This construct contains the following features: MCS, multiple cloning site for cloning the target protein in frame with the β-galΔω; GS Linker, (GGGGS)n; Hygro, hygromycin resistance gene; IRES, internal ribosome entry site; ColE1ori, origin of replication for growth in E. coli; 5'MoMuLV LTR and 3'MoMuLV LTR, viral promotor and polyadenylation signals from the Moloney Murine leukemia virus.

FIGURE 13B. Nucleotide sequence for pICAST OMN.

FIGURE 14. pICAST ALC β Arr2: Vector for expression of β -gal $\Delta\alpha$ as a C-terminal fusion to β -arrestin-2. The coding sequence of human β -arrestin-2 (Genebank Accession Number: NM_004313) was cloned in frame to β -gal $\Delta\alpha$ in a pICAST ALC vector.

FIGURE 15. pICAST OMC β Arr2: Vector for expression of β -gal $\Delta\omega$ as a C-terminal fusion to β -arrestin-2. The coding sequence of human β -arrestin-2 (Genebank Accession Number: NM_004313) was cloned in frame to β -gal $\Delta\omega$ in a pICAST OMC vector.

FIGURE 16. pICAST ALC β Arr1: Vector for expression of β -gal $\Delta\alpha$ as a C-terminal fusion to β -arrestin-1. The coding sequence of human β -arrestin-1 (Genebank Accession Number: NM_004041) was cloned in frame to β -gal $\Delta\alpha$ in a pICAST ALC vector.

FIGURE 17. pICAST OMC β Arr1: Vector for expression of β -gal $\Delta\omega$ as a C-terminal fusion to β -arrestin-1. The coding sequence of human β -arrestin-1 (Genebank Accession Number: NM_004041) was cloned in frame to β -gal $\Delta\omega$ in a pICAST OMC vector.

FIGURE 18. pICAST ALC β 2AR: Vector for expression of β -gal $\Delta\alpha$ as a C-terminal fusion to β 2 Adrenergic Receptor. The coding sequence of human β 2 Adrenergic Receptor

20

5

(Genebank Accession Number: NM_000024) was cloned in frame to β -gal $\Delta\alpha$ in a pICAST ALC vector.

FIGURE 19. pICAST OMC β 2AR: Vector for expression of β -gal $\Delta\omega$ as a C-terminal fusion β 2 Adrenergic Receptor. The coding sequence of human β 2 Adrenergic Receptor (Genebank Accession Number: NM_000024) was cloned in frame to β -gal $\Delta\omega$ in a pICAST OMC vector.

FIGURE 20. pICAST ALC A2aR: Vector for expression of β -gal $\Delta\alpha$ as a C-terminal fusion to Adenosine 2a Receptor. The coding sequence of human Adenosine 2a Receptor (Genebank Accession Number: NM_000675) was cloned in frame to β -gal $\Delta\alpha$ in a pICAST ALC vector.

FIGURE 21. pICAST OMC A2aR: Vector for expression of β -gal $\Delta\omega$ as a C-terminal fusion to Adenosine 2a Receptor. The coding sequence of human Adenosine 2a Receptor (Genebank Accession Number: NM_000675) was cloned in frame to β -gal $\Delta\omega$ in a pICAST OMC vector.

FIGURE 22. pICAST ALC D1: Vector for expression of β -gal $\Delta\alpha$ as a C-terminal fusion to Dopamine D1 Receptor. The coding sequence of human Dopamine D1 Receptor (Genebank Accession Number: X58987) was cloned in frame to β -gal $\Delta\alpha$ in a pICAST ALC vector.

FIGURE 23. A schematic depicting the method of the invention, which shows that two inactive mutants that become active when they interact.

20

5

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

All literature and patents cited in this disclosure are incorporated herein by reference.

The present invention provides a method to interrogate GPCR function and pathways. The G-protein-coupled superfamily continues to expand rapidly as new receptors are discovered through automated sequencing of cDNA libraries or genomic DNA. It is estimated that several thousand GPCRs may exist in the human genome, as many as 250 GPCRs have been cloned and only as few as 150 have been associated with ligands. The means by which these, or newly discovered orphan receptors, will be associated with their cognate ligands and physiological functions represents a major challenge to biological and biomedical research. The identification of an orphan receptor generally requires an individualized assay and a guess as to its function. The interrogation of a GPCR's signaling behavior by introducing a replacement receptor eliminates these prerequisites because it can be performed with and without prior knowledge of other signaling events. It is sensitive, rapid and easily performed and should be applicable to nearly all GPCRs because the majority of these receptors should desensitize by a common mechanism.

Various approaches have been used to monitor intracellular activity in response to a stimulant, <u>e.g.</u>, enzyme-linked immunosorbent assay (ELISA); Fluorescense Imaging Plate Reader assay (FLIPRTM, Molecular Devices Corp., Sunnyvale, CA); EVOscreenTM, EVOTECTM, Evotec Biosystems Gmbh, Hamburg, Germany; and techniques developed by CELLOMICSTM, Cellomics, Inc., Pittsburgh, PA.

Germino, F.J., et al., "Screening for in vivo protein-protein interactions." Proc. Natl. Acad. Sci., 90(3): 933-7 (1993), discloses an *in vivo* approach for the isolation of proteins interacting with a protein of interest.

20

5

Phizicky, E.M., et al., "Protein-protein interactions: methods for detection and analysis." Microbiol. Rev., 59(1): 94-123 (1995), discloses a review of biochemical, molecular biological and genetic methods used to study protein-protein interactions.

Offermanns, et al., " $G\alpha_{15}$ and $G\alpha_{16}$ Couple a Wide Variety of Receptors to Phospholipase C." J. Biol. Chem., 270(25):15175-80 (1995), discloses that $G\alpha_{15}$ and $G\alpha_{16}$ can be activated by a wide variety of G-protein-coupled receptors. The selective coupling of an activated receptor to a distinct pattern of G-proteins is regarded as an important requirement to achieve accurate signal transduction. <u>Id.</u>

Barak et al., "A β-arrestin/Green Fluorescent Protein Biosensor for Detecting G Protein-coupled Receptor Activation." J. Biol. Chem., 272(44):27497-500 (1997) and U.S. Patent No. 5,891,646, disclose the use of a β-arrestin/green fluorescent fusion protein (GFP) to monitor protein translocation upon stimulation of GPCR.

The present invention involves a method for monitoring protein-protein interactions in GPCR pathways as a complete assay using ICASTTM (Intercistronic Complementation Analysis Screening Technology as disclosed in pending U.S. patent application serial no. 053,164, filed April 1, 1998, the entire contents of which are incorporated herein by reference). This invention enables an array of assays, including GPCR binding assays, to be achieved directly within the cellular environment in a rapid, non-radioactive assay format amenable to high-throughput screening. Using existing technology, assays of this type are currently performed in a non-cellular environment and require the use of radioisotopes.

The present invention combined with Tropix ICASTTM and Advanced Discovery SciencesTM technologies, <u>e.g.</u>, ultra high-throughput screening, provide highly sensitive cell-based methods for interrogating GPCR pathways which are amendable to high-throughput

5

screening (HTS). These methods are an advancement over the invention disclosed in U.S. Patent 5,891,646, which relies on microscopic imaging of GPCR components as fusion with Green-fluorescent-protein. Imaging techniques are limited by low-throughput, lack of thorough quantification and low signal to noise ratios. Unlike yeast-based-2-hybrid assays used to monitor protein/protein interactions in high-throughput assays, the present invention is applicable to a variety of cells including mammalian cells, plant cells, protozoa cells such as E. coli and cells of invertebrate origin such as yeast, slime mold (*Dictyostelium*) and insects; detects interactions at the site of the receptor target or downstream target proteins rather than in the nucleus; and does not rely on indirect read-outs such as transcriptional activation. The present invention provides assays with greater physiological relevance and fewer false negatives.

Advanced Discovery SciencesTM is in the business of offering custom-developed screening assays optimized for individual assay requirements and validated for automation. These assays are designed by HTS experts to deliver superior assay performance. Advanced Discovery Sciences'TM custom assay development service encompasses the design, development, optimization and transfer of high performance screening assays. Advanced Discovery SciencesTM works to design new assays or convert existing assays to ultra-sensitive luminescent assays ready for the rigors of HTS. Among some of the technologies developed by Advanced Discovery SciencesTM are the cAMP-ScreenTM immunoassay system. This system provides ultrasensitive determination of cAMP levels in cell lysates. The cAMP-ScreenTM assay utilizes the high-sensitivity chemiluminescent alkaline phosphatase (AP) substrate CSPD® with Sapphire-IITM luminescence enhancer.

20

5

EXAMPLE:

GPCR activation can be measured through monitoring the binding of ligand-activated GPCR by an arrestin. In this assay system, a GPCR, e.g. β adrenergic receptor (β 2AR) and a β arrestin are co-expressed in the same cell as fusion proteins with β gal mutants. As illustrated in Figure 1, the β 2AR is expressed as a fusion protein with $\Delta\alpha$ form of β gal mutant (β 2ADR $\Delta\alpha$) and the β arrestin as a fusion protein with the $\Delta\omega$ mutant of β gal (β -Arr $\Delta\omega$). The two fusion proteins exist inside of a resting (or un-stimulated) cell in separate compartments, i.e. membrane for GPCR and cytosol for arrestin, and they can not form an active β galactosidase enzyme. When such a cell is treated with an agonist or a ligand, the ligand-occupied and activated receptor will become a high affinity binding site for Arrestin. The interaction between an activated β 2ADR $\Delta\alpha$ and β -Arr $\Delta\omega$ drives the β gal gal mutant complementation. The enzyme activity can be measured by using an enzyme substrate, which upon cleavage releases a product measurable by colorimetry, fluorescence, chemiluminescence (e.g. Tropix product GalScreenTM).

Experiment protocol-

- 1. In the first step, the expression vectors for $\beta 2ADR\Delta\alpha$ and $\beta Arr2\Delta\omega$ were engineered in selectable retroviral vectors pICAST ALC, as described in Figure 18 and pICAST OMC, as in Figure 15.
- 2. In the second step, the two expression constructs were transduced into either C2C12 myoblast cells, or other mammalian cell lines, such as COS-7, CHO, A431, HEK 293, and CHW. Following selection with antibiotic drugs, stable clones expressing both fusion

5

proteins at appropriate levels were selected.

3. In the last step, the cells expressing both $\beta 2ADR\Delta\alpha$ and $\beta Arr2\Delta\omega$ were tested for response by agonist/ligand stimulated β galactosidase activity. Triplicate samples of cells were plated at 10,000 cells in 100 microliter volume into a well of 96-well culture plate. Cells were cultured for 24 hours before assay. For agonist assay (Figure 3 and 4), cells were treated with variable concentrations of agonist, for example, (-) isoproterenol, procaterol, dobutamine, terbutiline or L-L-phenylephrine for 60 min at 37 C. The induced β galatosidase activity was measured by addition of Tropix GalScreenTM substrate (Applied Biosystems) and luminescence measured in a Tropix TR717TM luminometer (Applied Biosystems). For antagonist assay (Figure 5), cells were pre-incubated for 10 min in fresh medium without serum in the presence of ICI-118,551 or propranolol followed by addition of 10 micro molar (-) isoproterenol.

The assays of this invention, and their application and preparation have been described both generically, and by specific example. The examples are not intended as limiting. Other substituent identities, characteristics and assays will occur to those of ordinary skill in the art, without the exercise of inventive faculty. Such modifications remain within the scope of the invention, unless excluded by the express recitation of the claims advanced below.

20

5

WHAT IS CLAIMED IS:

- 1. A method of assessing the effect of a test condition on G-protein-coupled receptor (GPCR) pathway activity, comprising:
- a) providing a cell that expresses a GPCR as a fusion protein to one mutant form of reporter enzyme and an interacting protein partner as a fusion to another mutant form of enzyme;
 - b) exposing the cell to a ligand for said GPCR under said test condition; and
- c) monitoring activation of said GPCR by complementation of said reporter enzyme; wherein increased reporter enzyme activity in the cell compared to that which occurs in the absence of said test condition indicates increased GPCR interaction with its interacting protein partner compared to that which occurs in the absence of said test condition, and decreased reporter enzyme activity in the cell compared to that which occurs in the absence of said test condition indicates decreased GPCR interaction with its interacting protein partner compared to that which occurs in the absence of said test condition.
- 2. A method according to Claim 1, wherein the test condition is the presence in the cell of a kinase.
- 3. A method according to Claim 1, wherein the test condition is the presence in the cell of a G-protein.
- 4. A method according to Claim 1, wherein the test condition is the exposure of the cell to a compound selected from GPCR agonists and GPCR antagonists.
- 5. A method according to Claim 1, wherein the test condition is co-expression in the cell of a second receptor.
 - 6. A method according to Claim 5, wherein the second receptor is a GPCR receptor.

5

- 7. A method according to Claim 5, wherein homo-dimerization of GPCR is determined.
- 8. A method according to Claim 5, wherein hetero-dimerization of GPCR is determined.
- 9. A method for screening a β -arrestin protein or an unidentified arrestin or arrestinlike protein or fragment and mutant form thereof for the ability to bind to activated GPCRs, comprising:
 - a) providing a cell that:
 - i) expresses at least one GPCR as a fusion protein to a reporter enzyme; and
- ii) contains a conjugate comprising a test β -arrestin protein as a fusion protein with another reporter enzyme;
 - b) exposing the cell to a ligand for said at least one GPCR; and
- c) detecting enzymatic activity of the complemented reporter enzyme; $\text{wherein an increase in enzymatic activity in the cell indicates } \beta\text{-arrestin protein}$ binding to the activated GPCR.
- 10. A method for screening a test compound for G-protein-coupled receptor (GPCR) agonist activity, comprising:
- a) providing a cell that expresses a GPCR as a fusion protein to one mutant form of reporter enzyme and an arrestin protein as a fusion to another mutant form of enzyme;
 - b) exposing the cell to a test compound; and
 - c) detecting complementation of said reporter enzyme;

wherein increased reporter enzyme activity after exposure of the cell to the test compound indicates GPCR agonist activity of the test compound.

20

5

- 11. A method according to Claim 10, wherein the cell expresses a GPCR whose function is known.
- 12. A method according to Claim 10, wherein the cell expresses a GPCR whose function is unknown.
- 13. A method according to Claim 10, wherein the cell expresses an odorant or taste GPCR.
 - 14. A method according to Claim 10, wherein the cell expresses a GPCR a β -adrenergic GPCR.
 - 15. A method according to Claim 10, wherein the cell is selected from the group consisting of mammalian cells, cells of invertebrate origin, plant cells and protozoa cells.
 - 16. A method according to Claim 10, wherein the cell endogenously expresses a GPCR.
 - 17. A method according to Claim 10, wherein the cell has been transformed to express a GPCR not endogenously expressed by such a cell.
 - 18. A method of screening a test compound for G-protein-coupled receptor (GPCR) antagonist activity, comprising:
 - a) providing a cell that expresses a GPCR as a fusion protein to one mutant form of reporter enzyme and an arrestin protein as a fusion to another mutant form of enzyme;
 - b) exposing the cell to said test compound;
 - c) exposing the cell to an agonist for said GPCR; and
 - d) detecting complementation of said reporter enzyme;

where exposure to the agonist occurs at the same time as, or subsequent to, exposure to the test compound, and wherein decreased reporter enzyme activity after exposure of the

5

cell to the test compound indicates that the test compound is an antagonist for said GPCR.

- 19. A method of screening a cell for the presence of a G-protein-coupled receptor (GPCR) responsive to a GPCR agonist, comprising:
- a) providing a cell, said cell containing a conjugate comprising a β -arrestin protein as a fusion protein with a reporter enzyme;
 - b) exposing the cell to a GPCR agonist; and
 - c) detecting enzymatic activity of the reporter enzyme;

wherein an increase in enzymatic activity after exposure of the cell to the GPCR agonist indicates that the cell contains a GPCR responsive to said agonist.

- 20. A method of screening a plurality of cells for those cells which contain a G-protein-coupled receptor (GPCR) responsive to a GPCR agonist, comprising:
- a) providing a plurality of cells, said cells containing a conjugate comprising a β -arrestin protein as a fusion protein with a reporter enzyme;
 - b) exposing the cells to a GPCR agonist; and
 - c) detecting enzymatic activity of the reporter enzyme;

wherein an increase in enzymatic activity after exposure to the GPCR agonist indicates β -arrestin protein binding to a GPCR, thereby indicating that the cell contains a GPCR responsive to said GPCR agonist.

- 21. A method according to Claim 20, wherein the plurality of cells are contained in a tissue.
- 22. A method according to Claim 20, wherein the plurality of cells are contained in an organ.

5

- 23. A method according to Claim 20, wherein step (b) comprises exposing the cells to a plurality of GPCR agonists or ligand libraries.
- 24. A substrate having deposited thereon a plurality of cells, said cells expressing at least one GPCR as a fusion protein to one mutant form of reporter enzyme and an arrestin protein as a fusion to another mutant form of enzyme.
- 25. A substrate according to Claim 24, wherein the substrate contains an enzymelabile chemical group which, upon cleavage by the reporter enzyme, releases a product measurable by colorimetry, fluorescence or chemiluminescence.
- 26. A substrate according to Claim 24, wherein the substrate is made of a material selected from glass, plastic, ceramic, semiconductor, silica, fiber optic, diamond, biocompatible monomer and biocompatible polymer materials.
- 27. A method of detecting G-protein-coupled receptor (GPCR) pathway activity in a cell expressing at least one GPCR and containing β -arrestin protein as a fusion protein with a reporter enzyme; wherein said enzymatic activity indicates activation of the GPCR pathway.
- 28. A method according to Claim 27, where the cells are deposited on a substrate prior to detecting said enzymatic activity.
 - 29. A method according to Claim 27, wherein said cell is contained in a tissue.
 - 30. A method according to Claim 27, wherein said cell is contained in an organ.

ABSTRACT

Methods for detecting G-protein coupled receptor (GPCR) activity; methods of assaying GPCR activity; and methods of screening for GPCR ligands, G-protein-coupled receptor kinase (GRK) activity, and compounds that interact with components of the GPRC regulatory process are described.

Cellular Expression of β_2 AR- β gal $\Delta\alpha$ Fusion Protein in C2 Clones (measured by anti- β -gal ELISA)

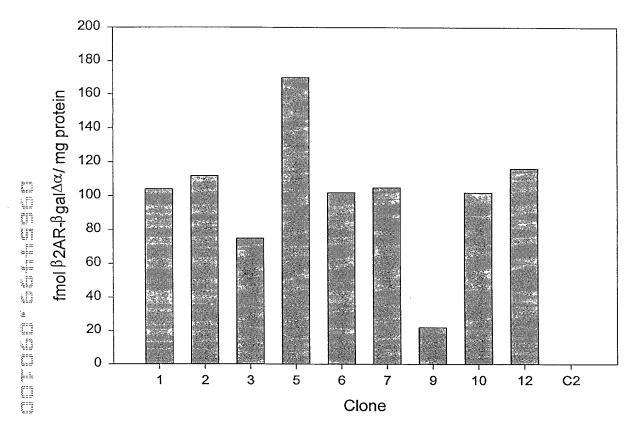


FIGURE 1A

Cellular expression of $\beta Arr2-\beta gal\Delta\omega$ fusion protein in C2 clones (measured by anti- β gal ELISA)

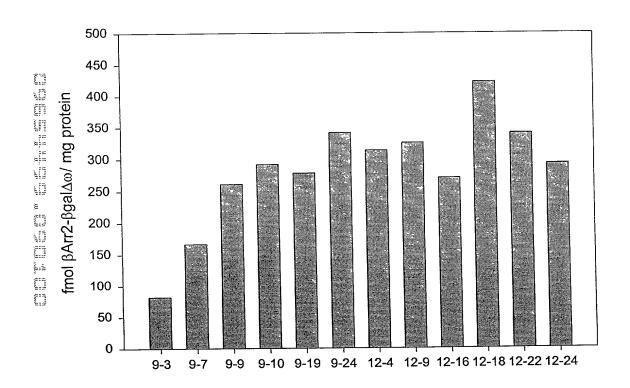


FIGURE 1B

Agonist Stimulated cAMP Response in C2 Cells Expressing β 2AR- β gal $\Delta\alpha$

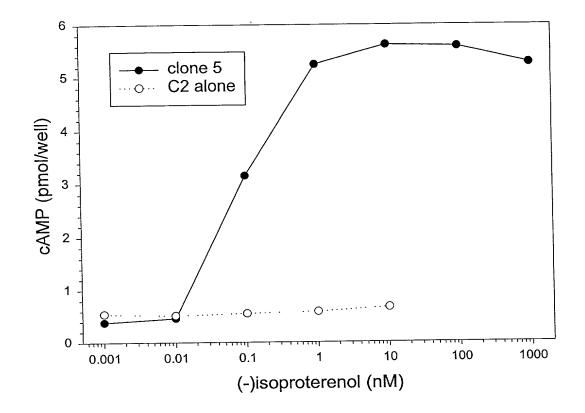
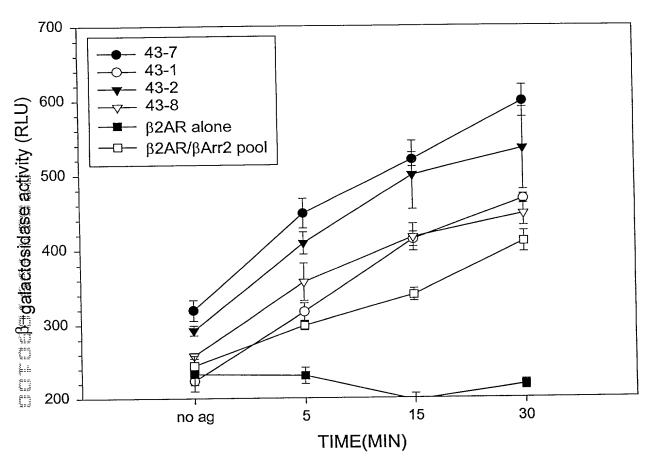
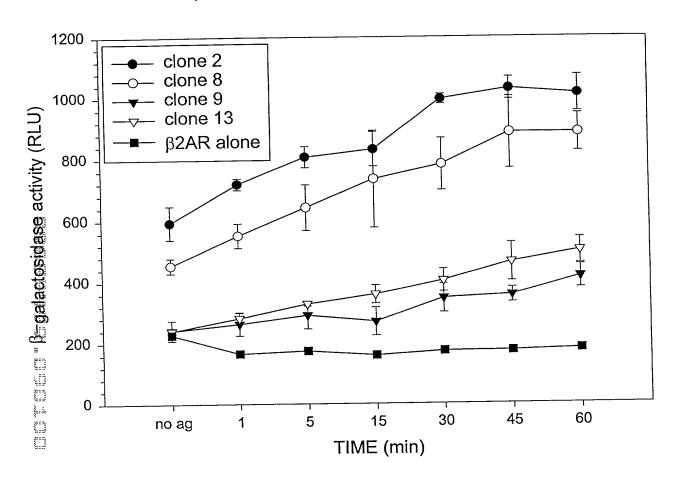


FIGURE 2

$\beta-$ galactosidase Complementation as a Measurement for $\beta 2AR-\beta gal\Delta\alpha$ interacting with $\beta Arrestin 2-\beta gal\Delta\omega$ upon agonist Stimulation



β –galactosidase Complementation as a Measurement for β 2AR- β gal $\Delta\alpha$ Interaction with β Arrestin1- β gal $\Delta\omega$ upon Agonist Stimulation



β –galactosidase Activity in Response to Agonist in C2 Cells Coexpressing β 2AR- β gal $\Delta\alpha$ and β Arrestin2- β gal $\Delta\omega$ Fusion Proteins

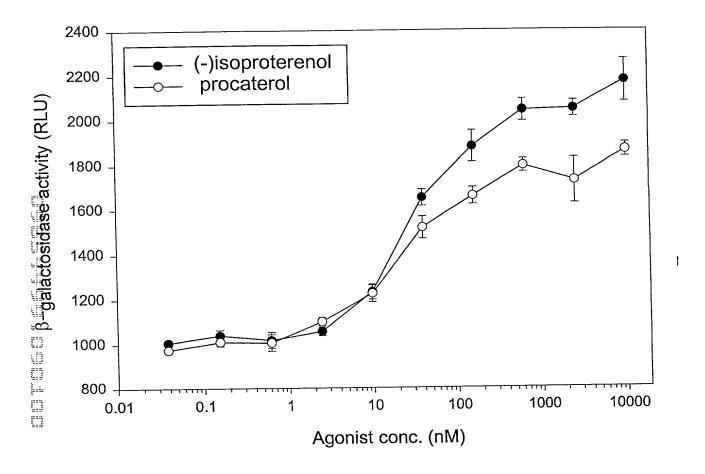
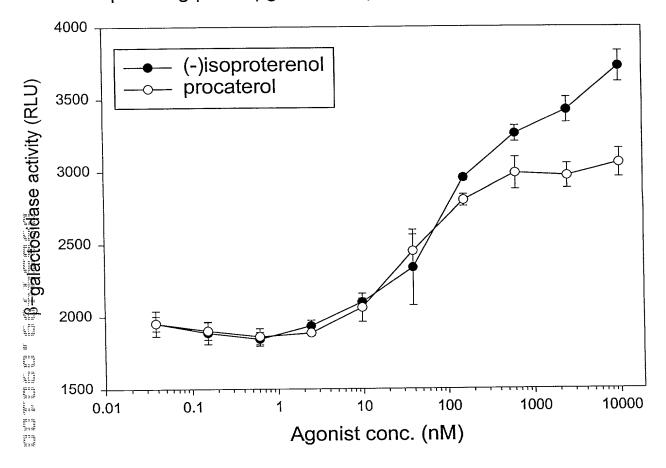


FIGURE 4A

$\beta-$ galactosidase Activity in Response to Agonist in C2 Cells Coexpressing $\beta2AR-\beta gal\Delta\alpha$ and $\beta Arrestin1-\beta gal\Delta\omega$ Fusion Proteins



Inhibition of β -galactosidase activity in C2 Cells Coexpressing β 2AR- β gal $\Delta\alpha$ and β Arrestin2- β gal $\Delta\omega$ Fusion Proteins

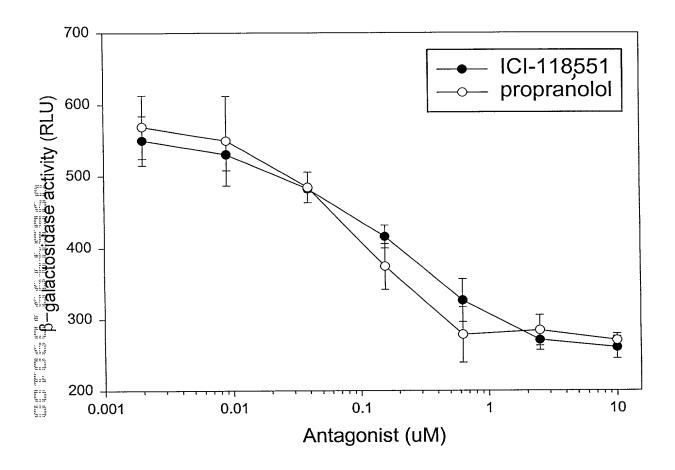
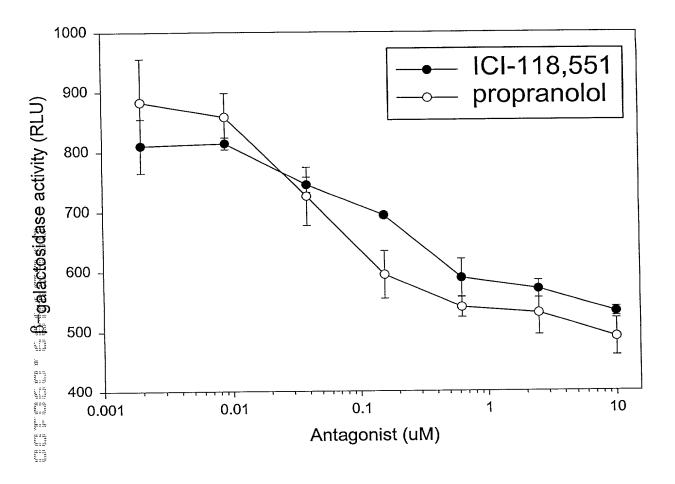


FIGURE 5A

Antagonist Inhibition of β –galactosidase Activity in C2 Cells Coexpressing β 2AR- β gal $\Delta\alpha$ and β Arrestin1- β gal $\Delta\omega$ Fusion Proteins



į

Figure 5B

Agonist Stimulated cAMP Response in Clones or Pools of C2 Cells Coexpressing A2aR- β gal $\Delta\alpha$ and β Arrestin1- β gal $\Delta\omega$ Fusion Proteins

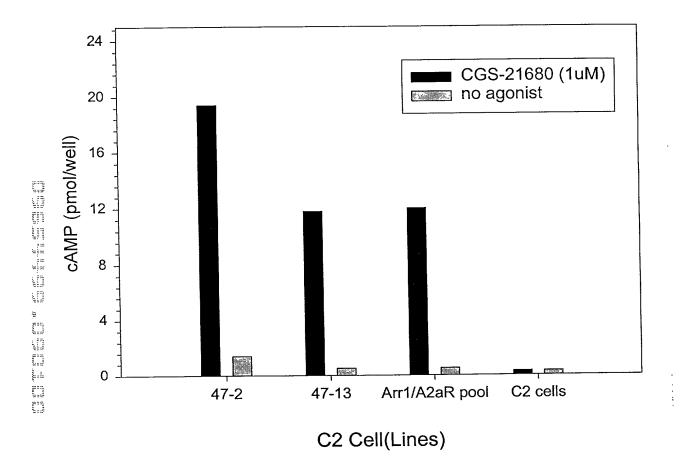


FIGURE 6

Agonist Stimulated cAMP Response in Clones or Pools of C2 Cells Expressing D1- β gal $\Delta\alpha$ and β Arrestin2- β gal $\Delta\omega$ Fusion Proteins

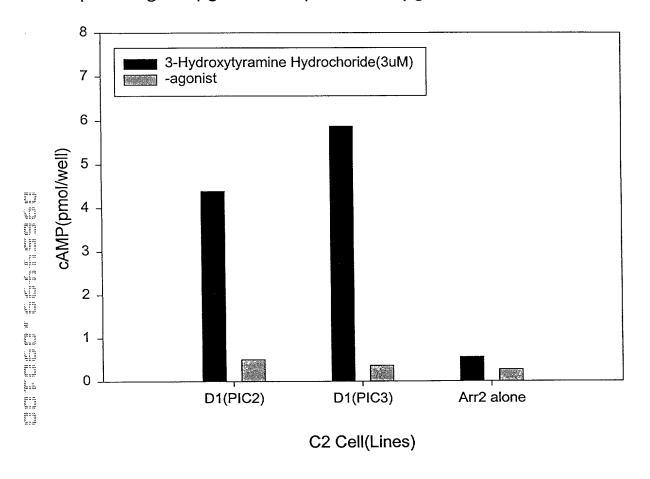


FIGURE 7

$β_2$ AR-βgalΔω and βarr2-βgalΔα Interaction in HEK293 Clones in Response to Isoproterenol Treatment (1 μM)

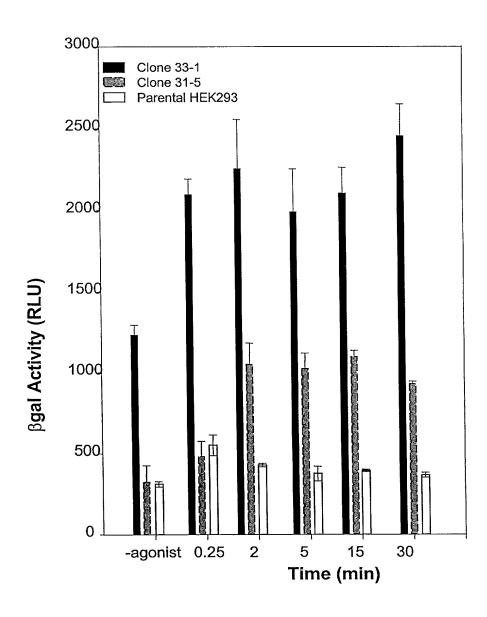
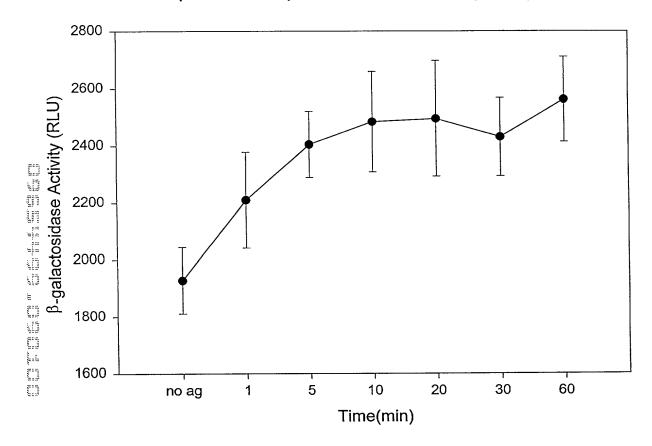
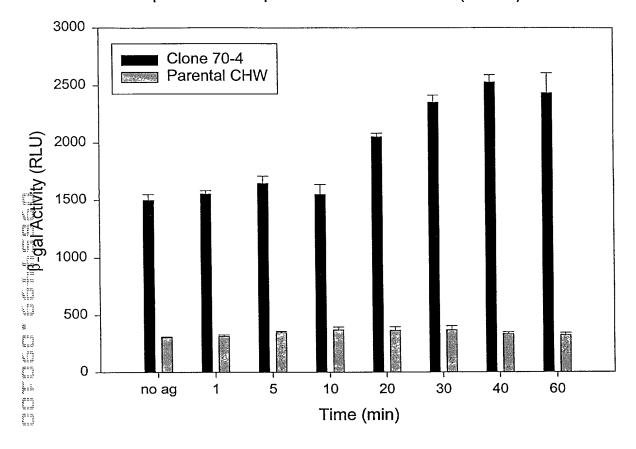


FIGURE 8A

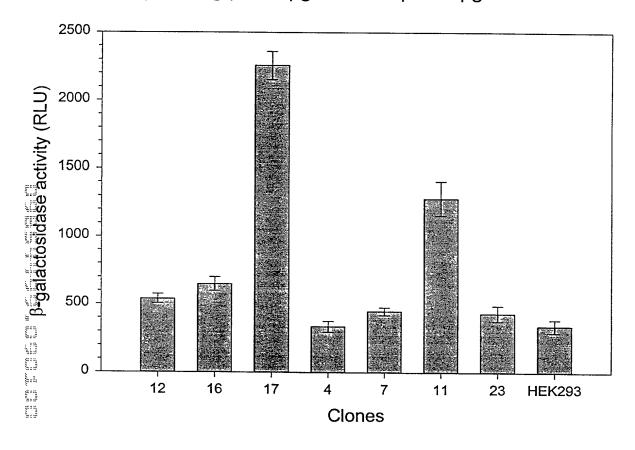
$\beta 2AR-\beta gal\Delta\alpha$ and $\beta Arr1-\beta gal\Delta$ Interaction in a CHO Pool in Response to Isoproterenol Treatment(10uM)



 $\beta 2AR-\beta gal\Delta\alpha$ and $\beta Arr2-\beta gal\Delta\omega$ Interaction in CHW Clone in Response to Isoproterenol Treatment (10uM)



 β –galactosidase Complementation as a Measurement for Adrenergic Receptor Homodimerization in HEK 293 Cells Coexpressing β 2AR- β gal $\Delta\alpha$ and β 2AR- β gal $\Delta\omega$.



Agonist Stimulated cAMP Response in HEK 293 Cells Coexpressing β 2AR- β gal $\Delta\alpha$ and β 2AR- β gal $\Delta\omega$

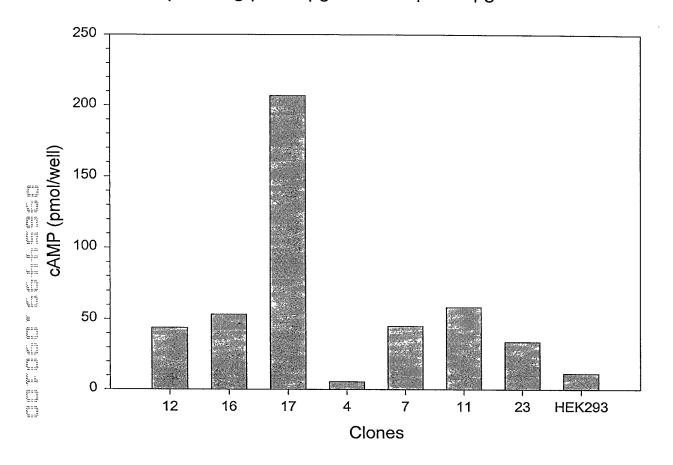


FIGURE 9B

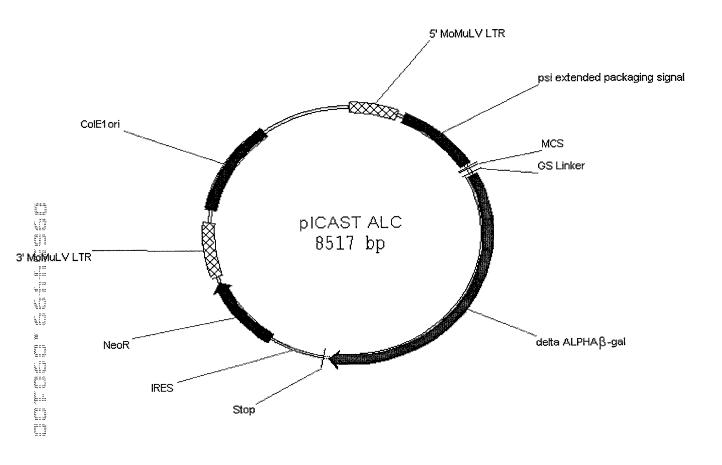


Figure 10A

1	CTGCAGCCTG GACGTCGGAC	AATATGGGCC TTATACCCGG				
51		GGGCCAAGAA CCCGGTTCTT				
101	GGATATCTGT CCTATAGACA	GGTAAGCAGT CCATTCGTCA				
151	GGTCCCCAGA CCAGGGGTCT	TGCGGTCCAG ACGCCAGGTC		AAAGATCTCT		
	GTTTCCAGGG CAAAGGTCCC	ACGGGGTTCC				
251	ATTGGTTAGT	GTTCGCTTCT CAAGCGAAGA	GCGAAGACAA	GCGCGCGAAG	TGCTCCCCGA ACGAGGGGCT	
		TCTCGGGTGT	TGGGGAGTGA	GCCCCGCGGT	CAGGAGGCTA	
351	TGACTGAGTC ACTGACTCAG	GCCCGGGTAC	CCGTGTATCC GGCACATAGG	AATAAACCCT TTATTTGGGA	CTTGCAGTTG GAACGTCAAC	
401	CATCCGACTT GTAGGCTGAA	GTGGTCTCGC CACCAGAGCG	TGTTCCTTGG ACAAGGAACC	GAGGGTCTCC CTCCCAGAGG	TCTGAGTGAT AGACTCACTA	
(m = 12)	ACTGATGGGC		AGAAAGTAAA	CCCCGAGCA	CCGGGATCGG GGCCCTAGCC	
501	GAGACCCCTG CTCTGGGGAC		TGGCTGGGTG	GTGGCCCTCC	CAAGCTGGCC CGTTCGACCGG	
	AGCAACTTAT				G ACTGATTTA C TGACTAAAAT	
601					CTGGCGGACC A GACCGCCTGG	
651					C CTGGGAGACG G GACCCTCTGC	
	AGGGTCCCTG	AAACCCCCGC	G CAAAAACACC	GGGCTGGAC	A GGAAGGGAGT I CCTTCCCTCA	
	CGATGTGGAA	A TCCGACCCC	G TCAGGATATO	G TGGTTCTGG	r aggagacgag a tcctctgctc	
801				T AAAAACGAA	T CGGTTTGGAA A GCCAAACCTT	
851	GGCTTCGGC	G CGCAGAACA	G ACGACGTCG	A TCGTTCTGT	G TTGTCTCTGT C AACAGAGACA	
	CTGACTGTGT GACTGACACA	r ttctgtatt A AAGACATAA	T GTCTGAAAA' A CAGACTTTT	T TAGGGCCAG A ATCCCGGTC	A CTGTTACCAC T GACAATGGTG	

951	TCCCTTAAGT TAGGGAATTCA					
1001	ACAACCAGTC (GGTAGATGTC		GTTGGGTTAC	GAAGACGAGA	
1051	GCAGAATGGC (CGTCTTACCG (GTTGGAAATT		GGCGCTCTGC	GCACCTTTAA	
1101	CCGAGACCTC A				GGACCGGGCG	
1151	ATGGACACCC ATACCTGTGGG	TCTGGTCCAG	GGGATGTAGC	ACTGGACCCT		
1201	TTTGACCCCC (AAACTGGGGG (
125 <u>1</u> - 11-	TCCTCTTCCT (GCAGAGAGGG			
1301	CCCCGCCTCG A	TAGGAGGGAA 	ATAGGTCGGG	AGTGAGGAAG	AGATCCGCGG	
		CGGGTAATTA	TGCTGAGTGA	TATCCCGCTA	AGCTTAGTCC	
_ 100_		GGCCTAGGAA 	TTAATTCGCG	TTAACCCTCC	ACCGCCATCG	
1451		GCGTGATTAC	GGATTCACTG	GCCGTCGTGG		•
- ==== +2	R P S	Q Q L I	R S L N	G E W		
1501	TCGCCCTTCC AGCGGGAAGG				CGCTTTGCCT GCGAAACGGA	
	W F P A GGTTTCCGGC	ACCAGAAGCG	GTGCCGGAAA	GCTGGCTGGA		
+2	P E A D					
1601	CCTGAGGCCG GGACTCCGGC			TTGACCGTCT	ACGTGCCAAT	
	D A P CGATGCGCCC			PIT	V N P	
1031					C CAGTTAGGCG	

	P F V P T E N P T G C Y S L T F N	
1701	CGTTTGTTCC CACGGAGAAT CCGACGGGTT GTTACTCGCT CACATTTAAT GCAAACAAGG GTGCCTCTTA GGCTGCCCAA CAATGAGCGA GTGTAAATTA	
+2	V D E S W L Q E G Q T R I I F D G	
1751	GTTGATGAAA GCTGGCTACA GGAAGGCCAG ACGCGAATTA TTTTTGATGG	
	CAACTACTTT CGACCGATGT CCTTCCGGTC TGCGCTTAAT AAAAACTACC	
+2	V N S A F H L W C N G R W V G Y	
1801	CGTTAACTCG GCGTTTCATC TGTGGTGCAA CGGGCGCTGG GTCGGTTACG GCAATTGAGC CGCAAAGTAG ACACCACGTT GCCCGCGACC CAGCCAATGC	
+2	G Q D S R L P S E F D L S A F L R	
1851	GCCAGGACAG TCGTTTGCCG TCTGAATTTG ACCTGAGCGC ATTTTTACGC CGGTCCTGTC AGCAAACGGC AGACTTAAAC TGGACTCGCG TAAAAATGCG	
1. +2	A G E N R L A V M V L R W S D G S	
190 1	GCCGGAGAAA ACCGCCTCGC GGTGATGGTG CTGCGCTGGA GTGACGGCAG	
	CGGCCTCTTT TGGCGGAGCG CCACTACCAC GACGCGACCT CACTGCCGTC	
+2		
1951	TTATCTGGAA GATCAGGATA TGTGGCGGAT GAGCGGCATT TTCCGTGACG AATAGACCTT CTAGTCCTAT ACACCGCCTA CTCGCCGTAA AAGGCACTGC	
	V S L L H K P T T Q I S D F H V A	
2001	TCTCGTTGCT GCATAAACCG ACTACACAAA TCAGCGATTT CCATGTTGCC	
	AGAGCAACGA CGTATTTGGC TGATGTGTTT AGTCGCTAAA GGTACAACGG	
+2	TRFNDDFSRAVLEAEVQ	
2051	ACTCGCTTTA ATGATGATTT CAGCCGCGCT GTACTGGAGG CTGAAGTTCA TGAGCGAAAT TACTACTAAA GTCGGCGCGA CATGACCTCC GACTTCAAGT	
+2	MCGELRDYLRVTVSLW	
2101	GATGTGCGGC GAGTTGCGTG ACTACCTACG GGTAACAGTT TCTTTATGGC	
	CTACACGCCG CTCAACGCAC TGATGGATGC CCATTGTCAA AGAAATACCG	
+2	Q G E T Q V A S G T A P F G G E I	
2151	AGGGTGAAAC GCAGGTCGCC AGCGGCACCG CGCCTTTCGG CGGTGAAATT TCCCACTTTG CGTCCAGCGG TCGCCGTGGC GCGGAAAGCC GCCACTTTAA	
 +:	2 I D E R G G Y A D R V T L R L N V	
2201	ATCGATGAGC GTGGTGGTTA TGCCGATCGC GTCACACTAC GTCTGAACGT	
	TAGCTACTCG CACCACCAAT ACGGCTAGCG CAGTGTGATG CAGACTTGCA	
	2 E N P K L W S A E I P N L Y R A	
2251	CGAAAACCCG AAACTGTGGA GCGCCGAAAT CCCGAATCTC TATCGTGCGG GCTTTTGGGC TTTGACACCT CGCGGCTTTA GGGCTTAGAG ATAGCACGCC	

+2	V V E L H T A D G T L I E A E A C	
2301	TGGTTGAACT GCACACCGCC GACGGCACGC TGATTGAAGC AGAAGCCTGC ACCAACTTGA CGTGTGGCGG CTGCCGTGCG ACTAACTTCG TCTTCGGACG	
+2	DVGFREVRIENGLLLN	
2351	GATGTCGGTT TCCGCGAGGT GCGGATTGAA AATGGTCTGC TGCTGCTGAA CTACAGCCAA AGGCGCTCCA CGCCTAACTT TTACCAGACG ACGACGACTT	
+2	G K P L L I R G V N R H E H H P	
2401	CGGCAAGCCG TTGCTGATTC GAGGCGTTAA CCGTCACGAG CATCATCCTC GCCGTTCGGC AACGACTAAG CTCCGCAATT GGCAGTGCTC GTAGTAGGAG	
+2	L H G Q V M D E Q T M V Q D I L L	
	TGCATGGTCA GGTCATGGAT GAGCAGACGA TGGTGCAGGA TATCCTGCTG ACGTACCAGT CCAGTACCTA CTCGTCTGCT ACCACGTCCT ATAGGACGAC	
[]- -[2	M K Q N N F N A V R C S H Y P N H	
250 <u>I</u>	ATGAAGCAGA ACAACTTTAA CGCCGTGCGC TGTTCGCATT ATCCGAACCA TACTTCGTCT TGTTGAAATT GCGGCACGCG ACAAGCGTAA TAGGCTTGGT	
	P L W Y T L C D R Y G L Y V V D	
2551		
- 755		
#2	E A N I E T H G M V P M N R L T D	
26044 -1 1 -	AAGCCAATAT TGAAACCCAC GGCATGGTGC CAATGAATCG TCTGACCGAT TTCGGTTATA ACTTTGGGTG CCGTACCACG GTTACTTAGC AGACTGGCTA	
[+ 2	DPRWLPAMSERVTRMVQ	
2651	GATCCGCGCT GGCTACCGGC GATGAGCGAA CGCGTAACGC GAATGGTGCA CTAGGCGCGA CCGATGGCCG CTACTCGCTT GCGCATTGCG CTTACCACGT	
+2	R D R N H P S V I I W S L G N E	
2701	GCGCGATCGT AATCACCCGA GTGTGATCAT CTGGTCGCTG GGGAATGAAT CGCGCTAGCA TTAGTGGGCT CACACTAGTA GACCAGCGAC CCCTTACTTA	
+2	SGHGANH DALYRWIKSV	
	CAGGCCACGG CGCTAATCAC GACGCGCTGT ATCGCTGGAT CAAATCTGTC GTCCGGTGCC GCGATTAGTG CTGCGCGACA TAGCGACCTA GTTTAGACAG	
	DPSRPVQYEGGGADTTA	
	GATCCTTCCC GCCCGGTGCA GTATGAAGGC GGCGGAGCCG ACACCACGGC CTAGGAAGGG CGGGCCACGT CATACTTCCG CCGCCTCGGC TGTGGTGCCG	
	T D I I C P M Y A R V D E D Q P	
2851	CACCGATATT ATTTGCCCGA TGTACGCGCG CGTGGATGAA GACCAGCCCT GTGGCTATAA TAAACGGGCT ACATGCGCGC GCACCTACTT CTGGTCGGGA	

+2	F P A V P K W S I K K W L S L P G
	TCCCGGCTGT GCCGAAATGG TCCATCAAAA AATGGCTTTC GCTACCTGGA
2901	AGGGCCGACA CGGCTTTACC AGGTAGTTTT TTACCGAAAG CGATGGACCT
+2	ETRPLIL CEYAHAM GNS
2951	GAGACGCGCC CGCTGATCCT TTGCGAATAC GCCCACGCGA TGGGTAACAG
	CTCTGCGCGG GCGACTAGGA AACGCTTATG CGGGTGCGCT ACCCATTGTC
+2	L G G F A K Y W Q A F R Q Y P R
	TCTTGGCGGT TTCGCTAAAT ACTGGCAGGC GTTTCGTCAG TATCCCCGTT AGAACCGCCA AAGCGATTTA TGACCGTCCG CAAAGCAGTC ATAGGGGCAA
	L Q G G F V W D W V D Q S L I K Y
3051	TACAGGGCGG CTTCGTCTGG GACTGGGTGG ATCAGTCGCT GATTAAATAT ATGTCCCGCC GAAGCAGACC CTGACCCACC TAGTCAGCGA CTAATTTATA
[] 43	DENGNPWSAYGGDFGDT
	GATGAAAACG GCAACCCGTG GTCGGCTTAC GGCGGTGATT TTGGCGATAC
31041	CTACTTTTGC CGTTGGGCAC CAGCCGAATG CCGCCACTAA AACCGCTATG
+2	PND RQFC MNG LVF ADR
3151	GCCGAACGAT CGCCAGTTCT GTATGAACGG TCTGGTCTTT GCCGACCGCA
a - 45-	CGGCTTGCTA GCGGTCAAGA CATACTTGCC AGACCAGAAA CGGCTGGCGT
	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$
ž ±	CGCCGCATCC AGCGCTGACG GAAGCAAAAC ACCAGCAGCA GTTTTTCCAG GCGGCGTAGG TCGCGACTGC CTTCGTTTTG TGGTCGTCGT CAAAAAAGGTC
- <u>1</u>	
+2	FRLSGQTIEVTSEYLFR
3251	TTCCGTTTAT CCGGGCAAAC CATCGAAGTG ACCAGCGAAT ACCTGTTCCG AAGGCAAATA GGCCCGTTTG GTAGCTTCAC TGGTCGCTTA TGGACAAGGC
+n	HSDNELLHWMVALDGK
3301	
2201	AGTATCGCTA TTGCTCGAGG ACGTGACCTA CCACCGCGAC CTACCATTCG
+2	PLASGEVPLDVAPQGKQ
3351	CGCTGGCAAG CGGTGAAGTG CCTCTGGATG TCGCTCCACA AGGTAAACAG
	GCGACCGTTC GCCACTTCAC GGAGACCTAC AGCGAGGTGT TCCATTTGTC
+2	LIEL PEL PQPESAG QLW
3401	TTGATTGAAC TGCCTGAACT ACCGCAGCCG GAGAGCGCCG GGCAACTCTG AACTAACTTG ACGGACTTGA TGGCGTCGGC CTCTCGCGGC CCGTTGAGAC
 +:	LTVRVVOPNATAWSEA
	GCTCACAGTA CGCGTAGTGC AACCGAACGC GACCGCATGG TCAGAAGCCG
2421	CGAGTGTCAT GCGCATCACG TTGGCTTGCG CTGGCGTACC AGTCTTCGGC

+2	G I	н І	S	Α	. V	J (2	Q	W	R	.]	L	Α	E	2 1	N	L	,	s	V
3501	GGC	ACAT	CAG	CGC	 CT(GC.	AG	CAC	GTG(GCG	TC	TG	GC(GGF	AA.	 A (CCT	CA	 GTG	TG
	CCG'	IGTA 	.GTC 	GCG	GA(CCG'	ГС 	GT(CAC(CGC 	:AG 	AC 	CG(CCI	TT' 	T (GGA 	.GT 	CAC	:AC
+2	T	L	P 2	A 	A 	s 	H 		A	I 	P	H 		ն 	T 	T 		s 	E 	M
3551			CCG GGC																	
+2	D	F	С	I	E	L	G	; 	N	K	R		W	Q	F		N	R	. Ç	<u> </u>
3601			TGC																	
+2	s		L																	
3651	CAG GTC	GCTI		TTC	CAC	AGA	TG AC	TG AC	GAT CTA	TG(GCG CGC	AT TA	AA' TT	AA.	AAC TTG	A T	ACI TG	GC ACG	TGA SACT	ACG IGC
 +2			R				Т			Α	Р	I		D	N	D)	I	G	V
370]	CCG	CTG	CGCG GCGC	ATO	CAG	TTC	AC	CC	GTG CAC	CA(CCG GGC	CI G	rgg Acc	AT.	AAC TTC	G	AC	$\Gamma T A$	GGG	
+2	 S	 E	A	T	- - R		 : I	- - -				. - -				· }	R	W	 V I	 K
375 <u>1</u>			AGCG ICGC																	
- []- [+2	- А		 G H																	
3801	CGG	CGG		TT	ACC	AGC	GCC CGG	GA CT	AGC	CAG	CGI GC <i>I</i>	TO A A	GTI CAA	GC CG	AGT TC	rg AC	CA(CGC	GCA	GAT
- 4.3 - £22	т			D	 A				. .									Q	 Н	 -
3851	ACA			AT	 GCG	GT	 GCT	 GP	 ATT <i>I</i>		ACC	 C G	CTC	CAC	GC	 3T	GG	CA	GCA CGT	 TCA AGT
+2				 T.	 4		 [s	 R	 K							D		 G	 S
3901	GG(GGAA		 C TT	 TTA	 'AT'	 rca	 . GO	CCG	 GAA	 AA(c c	 TAC		GA	 TT	GA	TG	 GTA	GTG
	 ? G																			- -
3951	GT	CAAA	TGG	C GA	TTZ		 GTT		ATG	TTG	AA:	 G T	GG(CGI	 AGC	 GA	 ТА	 CA	 .CCG	CAT
	CA	GTTI	'ACC	G CI	'AA'	ľGG	CAA	C:	TAC.	AAC	CTT	C A	CC	GC1	ľCG	CT	PΑ	GT	GGC	GTP
	2 P	A	R	I 	G	L	N	1 	C 	Q	L		Α	Q 	V 		A 	E		7
4001			GCC'											GT(CCA	TC	GI	'CT	'CGC	CCCI
	 : 2																			 T
4051		TGA	CCGA	G C	CTA	ATC	CCC	G G	CGT	TC:	rtt	T C	SAT	AG	GGC	TG	G	CGG	CAA	rga(

+2 A A C F D R W D L P L S D M Y T P
4101 CCGCCTGTTT TGACCGCTGG GATCTGCCAT TGTCAGACAT GTATACCCCG GGCGGACAAA ACTGGCGACC CTAGACGGTA ACAGTCTGTA CATATGGGGC
+2 Y V F P S E N G L R C G T R E L N
4151 TACGTCTTCC CGAGCGAAAA CGGTCTGCGC TGCGGGACGC GCGAATTGAA ATGCAGAAGG GCTCGCTTTT GCCAGACGCG ACGCCCTGCG CGCTTAACTT
+2 Y G P H Q W R G D F Q F N I S R
4201 TTATGGCCCA CACCAGTGGC GCGGCGACTT CCAGTTCAAC ATCAGCCGCT AATACCGGGT GTGGTCACCG CGCCGCTGAA GGTCAAGTTG TAGTCGGCGA
+2 Y S Q Q L M E T S H R H L L H A
4251 ACAGTCAACA GCAACTGATG GAAACCAGCC ATCGCCATCT GCTGCACGCG TGTCAGTTGT CGTTGACTAC CTTTGGTCGG TAGCGGTAGA CGACGTGCGC
ta E G T W L N I D G F H M G I G G
430 GAAGAAGGCA CATGGCTGAA TATCGACGGT TTCCATATGG GGATTGGTGG
CTTCTTCCGT GTACCGACTT ATAGCTGCCA AAGGTATACC CCTAACCACC
+2 D D S W S P S V S A E F Q L S A
4351 CGACGACTCC TGGAGCCCGT CAGTATCGGC GGAATTCCAG CTGAGCGCCG GCTGCTGAGG ACCTCGGGCA GTCATAGCCG CCTTAAGGTC GACTCGCGGC
+2 G R Y H Y Q L V W C Q K R S D Y K
4401 GTCGCTACCA TTACCAGTTG GTCTGGTGTC AAAAAAGATC TGACTATAAA CAGCGATGGT AATGGTCAAC CAGACCACAG TTTTTTCTAG ACTGATATTT
- El-
手2 D E D L D H H H H H R >
4451 GATGAGGACC TCGACCATCA TCATCATCAT CACCGGTAAT AATAGGTAGA CTACTCCTGG AGCTGGTAGT AGTAGTAGTA GTGGCCATTA TTATCCATCT
4501 TAAGTGACTG ATTAGATGCA TTGATCCCTC GACCAATTCC GGTTATTTTC
ATTCACTGAC TAATCTACGT AACTAGGGAG CTGGTTAAGG CCAATAAAAG
4551 CACCATATTG CCGTCTTTTG GCAATGTGAG GGCCCGGAAA CCTGGCCCTG GTGGTATAAC GGCAGAAAAC CGTTACACTC CCGGGCCTTT GGACCGGGAC
4601 TCTTCTTGAC GAGCATTCCT AGGGGTCTTT CCCCTCTCGC CAAAGGAATG
AGAAGAACTG CTCGTAAGGA TCCCCAGAAA GGGGAGAGCG GTTTCCTTAC
4651 CAAGGTCTGT TGAATGTCGT GAAGGAAGCA GTTCCTCTGG AAGCTTCTTG GTTCCAGACA ACTTACAGCA CTTCCTTCGT CAAGGAGACC TTCGAAGAAC
4701 AAGACAAACA ACGTCTGTAG CGACCCTTTG CAGGCAGCGG AACCCCCCAC
TTCTGTTTGT TGCAGACATC GCTGGGAAAC GTCCGTCGCC TTGGGGGGGTG
4751 CTGGCGACAG GTGCCTCTGC GGCCAAAAGC CACGTGTATA AGATACACCT
GACCGCTGTC CACGGAGACG CCGGTTTTCG GTGCACATAT TCTATGTGGA

4801	GCAAAGGCGG CGTTTCCGCC					
4851	AAGAGTCAAA TTCTCAGTTT					
						
4901	CCCAGAAGGT GGGTCTTCCA					
4951	TGCTTTACAT ACGAAATGTA					
						
5001	CACGGGGACG GTGCCCCTGC					
5051	AACAAGATGG TTGTTCTACC					•
2 2	TTCGGCTATG AAGCCGATAC				ATGCCGCCGT TACGGCGGCA	
##						
g cod dan g con		TCAGCGCAGG AGTCGCGTCC	CCGCGGGCCA	AGAAAAACAG	TTCTGGCTGG	
					COMPROCES	
5291 11	TGTCCGGTGC ACAGGCCACG	GGACTTACTT	GACGTCCTGC	TCCGTCGCGC	GCTATCGTGG CGATAGCACC	
5251					TTGTCACTGA AACAGTGACT	
	·					
4.4	AGCGGGAAGG TCGCCCTTCC				CAGGATCTCC CGTCCTAGAGG	
2.4						
5351					GGCTGATGCA CCGACTACGT	
E 4 O 1	ATGCGGCGGC	mccama cccm	mca mcccccm	አ ሮሮሞሮሮሮሮሊ ባ	י ייירכארראררא	
5401	TACGCCGCCG	ACGTATGCGA	ACTAGGCCGA	TGGACGGGT	A AGCTGGTGGT	
5451	AGCGAAACAT TCGCTTTGTA				A GCCGGTCTTG CGGCCAGAAC	
5501	AGCTAGTCCT	ACTAGACCTG	CTTCTCGTAG	TCCCCGAGC	C GCCAGCCGAA G CGGTCGGCTT	
5551	CTGTTCGCCA	GGCTCAAGGC	GCGCATGCCC	GACGGCGAG	G ATCTCGTCGT C TAGAGCAGCA	
	GACCCATGGC CTGGGTACCG	GATGCCTGCT CTACGGACGA	TGCCGAATATA ACGGCTTATA	CATGGTGGA GTACCACCT	A AATGGCCGCT T TTACCGGCGA	
5651	TTTCTGGATT AAAGACCTAA	CATCGACTGA GTAGCTGACA	GGCCGGCTGG CCGGCCGACG	G GTGTGGCGG C CACACCGCC	A CCGCTATCAG T GGCGATAGTC	
5701	GACATAGCGT	TGGCTACCC	G TGATATTGC	r gaagagctt	G GCGGCGAATG C CGCCGCTTAC	

5751	GGCTGACCGC TTCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTCGCAGC CCGACTGGCG AAGGAGCACG AAATGCCATA GCGGCGAGGG CTAAGCGTCG	
5801	GCATCGCCTT CTATCGCCTT CTTGACGAGT TCTTCTGAGC GGGACTCTGG CGTAGCGGAA GATAGCGGAA GAACTGCTCA AGAAGACTCG CCCTGAGACC	
5851	GGTTCGCATC GATAAAATAA AAGATTTTAT TTAGTCTCCA GAAAAAGGGG CCAAGCGTAG CTATTTTATT TTCTAAAATA AATCAGAGGT CTTTTTCCCC	
5901	GGAATGAAAG ACCCCACCTG TAGGTTTGGC AAGCTAGCTT AAGTAACGCC CCTTACTTTC TGGGGTGGAC ATCCAAACCG TTCGATCGAA TTCATTGCGG	
5951	ATTTTGCAAG GCATGGAAAA ATACATAACT GAGAATAGAG AAGTTCAGAT TAAAACGTTC CGTACCTTTT TATGTATTGA CTCTTATCTC TTCAAGTCTA	_
6001	CAAGGTCAGG AACAGATGGA ACAGCTGAAT ATGGGCCAAA CAGGATATCT GTTCCAGTCC TTGTCTACCT TGTCGACTTA TACCCGGTTT GTCCTATAGA	_
		-
6 0 51	GTGGTAAGCA GTTCCTGCCC CGGCTCAGGG CCAAGAACAG ATGGAACAGC CACCATTCGT CAAGGACGG GCCGAGTCCC GGTTCTTGTC TACCTTGTCG	_
6101	TGAATATGGG CCAAACAGGA TATCTGTGGT AAGCAGTTCC TGCCCCGGCT ACTTATACCC GGTTTGTCCT ATAGACACCA TTCGTCAAGG ACGGGGCCGA	_
4151 6151	CAGGGCCAAG AACAGATGGT CCCCAGATGC GGTCCAGCCC TCAGCAGTTT GTCCCGGTTC TTGTCTACCA GGGGTCTACG CCAGGTCGGG AGTCGTCAAA	_
6201	CTAGAGAACC ATCAGATGTT TCCAGGGTGC CCCAAGGACC TGAAATGACC GATCTCTTGG TAGTCTACAA AGGTCCCACG GGGTTCCTGG ACTTTACTGG	_
6251	CTGTGCCTTA TTTGAACTAA CCAATCAGTT CGCTTCTCGC TTCTGTTCGC GACACGGAAT AAACTTGATT GGTTAGTCAA GCGAAGAGCG AAGACAAGCG	
[3		-
63 01	GCGCTTCTGC TCCCCGAGCT CAATAAAAGA GCCCACAACC CCTCACTCGG CGCGAAGACG AGGGGCTCGA GTTATTTTCT CGGGTGTTGG GGAGTGAGCC	. -
6351	GGCGCCAGTC CTCCGATTGA CTGAGTCGCC CGGGTACCCG TGTATCCAAT CCGCGGTCAG GAGGCTAACT GACTCAGCGG GCCCATGGGC ACATAGGTTA	
6401	AAACCCTCTT GCAGTTGCAT CCGACTTGTG GTCTCGCTGT TCCTTGGGAG TTTGGGAGAA CGTCAACGTA GGCTGAACAC CAGAGCGACA AGGAACCCTC	
6451	GGTCTCCTCT GAGTGATTGA CTACCCGTCA GCGGGGGTCT TTCATTCATG CCAGAGGAGA CTCACTAACT GATGGGCAGT CGCCCCCAGA AAGTAAGTAC	- -
	CAGCATGTAT CAAAATTAAT TTGGTTTTTT TTCTTAAGTA TTTACATTAA GTCGTACATA GTTTTAATTA AACCAAAAAA AAGAATTCAT AAATGTAATT	
6552	ATGGCCATAG TTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TACCGGTATC AACGTAATTA CTTAGCCGGT TGCGCGCCCC TCTCCGCCAA	
	TGCGTATTGG CGCTCTTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG ACGCATAACC GCGAGAAGGC GAAGGAGCGA GTGACTGAGC GACGCGAGCC	
665	TCGTTCGGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG AGCAAGCCGA CGCCGCTCGC CATAGTCGAG TGAGTTTCCG CCATTATGCC	

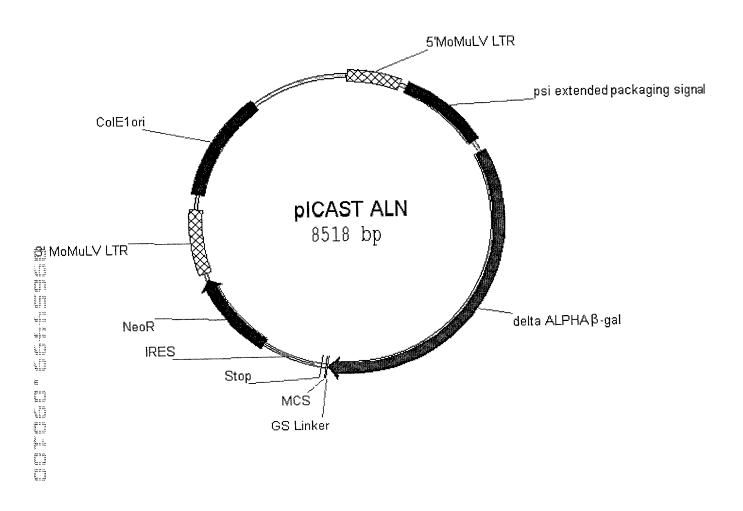


Figure 11A

1	CTGCAGCCTG	AATATGGGCC .	AAACAGGATA	TCTGTGGTAA	GCAGTTCCTG	
	GACGTCGGAC	TTATACCCGG	TTTGTCCTAT	AGACACCATT	CGTCAAGGAC	
51	CCCCGGCTCA GGGGCCGAGT	GGGCCAAGAA CCCGGTTCTT	CAGATGGAAC GTCTACCTTG	AGCTGAATAT TCGACTTATA	CCCGGTTTGT	
101	GGATATCTGT CCTATAGACA	GGTAAGCAGT CCATTCGTCA	TCCTGCCCCG AGGACGGGGC	GCTCAGGGCC CGAGTCCCGG	TTCTTGTCTA	
151	CCAGGGGTCT	TGCGGTCCAG ACGCCAGGTC	GGGAGTCGTC	AAAGATCTCT	TGGTAGTCTA	
201	GTTTCCAGGG	TGCCCCAAGG ACGGGGTTCC	ACCTGAAATG	ACCCTGTGCC	TTATTTGAAC AATAAACTTG	
251	ATTGGTTAGT	GTTCGCTTCT CAAGCGAAGA	GCGAAGACAA	CGCGCGCTTC	ACGAGGGGC'1'	
701		AGAGCCCACA		CGGGGCGCCA	GTCCTCCGAT	
301 []	CGAGTTATTT	TCTCGGGTGT	TGGGGAGTGA	GCCCGCGGT	CAGGAGGCTA	
351	TGACTGAGTC	GCCCGGGTAC	CCGTGTATCC	AATAAACCCI	CTTGCAGTTG	
	ACTGACTCAG	CGGGCCCATG	GGCACATAGG	TTATTTGGGA	GAACGTCAAC	
401	CATCCGACTT	GTGGTCTCGC	TGTTCCTTGG	GAGGGTCTCC	TCTGAGTGAT	
VI		CACCAGAGCG			G AGACTCACTA	
4 51	TGACTACCCG	TCAGCGGGGG	TCTTTCATTI	GGGGGCTCG	CCGGGATCGG	
%.: 121	ACTGATGGGC	AGTCGCCCCC	AGAAAGTAAA	CCCCGAGC	A GGCCCTAGCC	
501	GAGACCCCTG	CCCAGGGACC	ACCGACCCAC	CACCGGGAG	G CAAGCTGGCC	
	CTCTGGGGAC	GGGTCCCTGG	TGGCTGGGT	GTGGCCCTC	C GTTCGACCGG	
551					G ACTGATTTTA	•-
"ada∪ I	TCGTTGAATA	A GACACAGACA	GGCTAACAG	A TCACAGATA	C TGACTAAAAT	
		 _	TTAGCTAAC	r agctctgta	T CTGGCGGACC	• •
601	ACGCGGACG(C AGCCATGATO	AATCGATTG	A TCGAGACAT	A GACCGCCTGG	
651	CGTGGTGGA GCACCACCT	A CTGACGAGT' T GACTGCTCA <i>I</i>	r CTGAACACC A GACTTGTGG	G CCGGCGTTG	C CTGGGAGACG G GACCCTCTGC	·
701	TCCCAGGGA AGGGTCCCT	C TTTGGGGGCG G AAACCCCCGG	C GTTTTTGTG G CAAAAACAC	G CCCGACCTG C GGGCTGGAC	A GGAAGGGAGT CCTTCCCTCA	
		- =======				
751	GCTACACCT	T AGGCTGGGG	C AGTCCTATA	C ACCAAGACO	T AGGAGACGAG A TCCTCTGCTC	
801	ААССТАААА	C AGTTCCCGC	C TCCGTCTGF	A TTTTTGCT	T CGGTTTGGAA	
		G TCAAGGGCG			AA GCCAAACCTT	
851	CCGAAGCCG	C GCGTCTTGT	C TGCTGCAG	CA TCGTTCTG	rg ttgtctctgt AC AACAGAGACA	-
901	CTGACTGT	 ST TTCTGTATI	T GTCTGAAA	AT TAGGGCCA	GA CTGTTACCAC	- .
	GACTGACA	CA AAGACATAA	A CAGACTTT	TA ATCCCGGT	CT GACAATGGTG	

951	TCCCTTAAGT AGGGAATTCA	TTGACCTTAG AACTGGAATC	GTAACTGGAA CATTGACCTT	AGATGTCGAG TCTACAGCTC	CGGCTCGCTC GCCGAGCGAG	
1001	ACAACCAGTC TGTTGGTCAG	GGTAGATGTC CCATCTACAG	AAGAAGAGAC TTCTTCTCTG	GTTGGGTTAC CAACCCAATG	CTTCTGCTCT GAAGACGAGA	
1051	GCAGAATGGC CGTCTTACCG	CAACCTTTAA GTTGGAAATT	CGTCGGATGG GCAGCCTACC	CCGCGAGACG GGCGCTCTGC	GCACCTTTAA CGTGGAAATT	
1101	CCGAGACCTC GGCTCTGGAG	ATCACCCAGG TAGTGGGTCC	TTAAGATCAA AATTCTAGTT	GGTCTTTTCA CCAGAAAAGT	CCTGGCCCGC GGACCGGGCG	
1151	ATGGACACCC TACCTGTGGG	AGACCAGGTC TCTGGTCCAG	CCCTACATCG GGGATGTAGC	TGACCTGGGA ACTGGACCCT	AGCCTTGGCT TCGGAACCGA	
1201	TTTGACCCCC AAACTGGGGG	CTCCCTGGGT GAGGGACCCA	CAAGCCCTTT GTTCGGGAAA	GTACACCCTA CATGTGGGAT	AGCCTCCGCC TCGGAGGCGG	
1251 13	TCCTCTTCCT AGGAGAAGGA	CCATCCGCCC GGTAGGCGGG	CGTCTCTCCC GCAGAGAGGG	CCTTGAACCT	CCTCGTTCGA GGAGCAAGCT	
1 901	CCCCGCCTCG GGGGCGGAGC	ATCCTCCCTT TAGGAGGGAA	TATCCAGCCC ATAGGTCGGG	TCACTCCTTC AGTGAGGAAG	TCTAGGCGCC AGATCCGCGG	
1351 13	GGCCGCTCTA CCGGCGAGAT	GCCCATTAAT CGGGTAATTA	ACGACTCACT TGCTGAGTGA	ATAGGGCGAT	TCGAACACCA AGCTTGTGGT	
1401	TGCACCATCA ACGTGGTAGT	TCATCATCAC AGTAGTAGTG	GTCGACTATA CAGCTGATAT	AAGATGAGGA TTCTACTCCT	CCTCGAGATG GGAGCTCTAC	
	GGCGTGATTA CCGCACTAAT	CGGATTCACT GCCTAAGTGA	GGCCGTCGTG	G GCCCGCACCC C CGGGCGTGGC	ATCGCCCTTC TAGCGGGAAG	
	CCAACAGTTA GGTTGTCAAT	CGCAGCCTGA GCGTCGGACT	ATGGCGAAT(G GCGCTTTGCGCCCCGCGAAACGG	TGGTTTCCGG ACCAAAGGCC	
1551	CACCAGAAGC GTGGTCTTCG	GGTGCCGGAA	A AGCTGGCTGG	G AGTGCGATC	T TCCTGAGGCC A AGGACTCCGG	
1601	GATACTGTCG CTATGACAGC	TCGTCCCCTC	C AAACTGGCA	G ATGCACGGT C TACGTGCCA	T ACGATGCGCC A TGCTACGCGG	
	GTAGATGTGG	TTGCACTGG	A TAGGGTAAT	G CCAGTTAGG	G CCGTTTGTTC C GGCAAACAAG	
	CCACGGAGA!	A TCCGACGGG' AGGCTGCCC	T TGTTACTCG A ACAATGAGC	C TCACATTTA G AGTGTAAAT	A TGTTGATGAA T ACAACTACTT	
	TCGACCGAT	G TCCTTCCGG	T CTGCGCTTA	A TAAAAACTA	G GCGTTAACTC C CGCAATTGAG	
1801	GGCGTTTCA'	T CTGTGGTGC A GACACCACG	A ACGGGCGCT T TGCCCGCGA	G GGTCGGTTA	AC GGCCAGGACA	
1851	GTCGTTTGC CAGCAAACG	C GTCTGAATT G CAGACTTAA	T GACCTGAGO A CTGGACTCO	CG CATTTTTAC	CG CGCCGGAGAA GC GCGGCCTCTT	

1901	AACCGCCTCG C	GGTGATGGT	GCTGCGCTGG	AGTGACGGCA	GTTATCTGGA	
	TTGGCGGAGC G					
1951	AGATCAGGAT A	ACACCGCCT	TGAGCGGCAT ACTCGCCGTA	AAAGGCACTG	CAGAGCAACG	
2001	TGCATAAACC G ACGTATTTGG G	SACTACACAA	ATCAGCGATT	TCCATGTTGC	GTGAGCGAAA	
2051	AATGATGATT T	AGTCGGCGCG	ACATGACCTC	CGACTTCAAG	TCTACACGCC	
2101	CGAGTTGCGT (GCTCAACGCA (GACTACCTAC	GGGTAACAGT	TTCTTTATGG	CAGGGTGAAA	
2151	CGCAGGTCGC (CAGCGGCACC GTCGCCGTGG	GCGCCTTTCG CGCGGAAAGC	GCGGTGAAAT CGCCACTTTA	ATAGCTACTC	
2 2 01	CGTGGTGGTT . GCACCACCAA	ATGCCGATCG TACGGCTAGC	CGTCACACTA GCAGTGTGAT	GCAGACTTGC	AGCTTTTGGG	
2251	GAAACTGTGG CTTTGACACC	AGCGCCGAAA TCGCGGCTTT	TCCCGAATCT AGGGCTTAGA	CTATCGTGCG GATAGCACGC	GTGGTTGAAC CACCAACTTG	
:== .						
2301		GCTGCCGTGC	GACTAACTTC	CAGAAGCCTG CGTCTTCGGAC	GCTACAGCCA	
2351	TTCCGCGAGG	TGCGGATTGA	AAATGGTCT	CTGCTGCTG	A ACGGCAAGCC TGCCGTTCGG	
	GTTGCTGATT CAACGACTAA	GCTCCGCAAT	TGGCAGTGC	r CGTAGTAGG	A GACGTACCAG	
2451	ACCTCATGGA	TGAGCAGACG	ATGGTGCAG	ATATCCTGC		
2501	AACAACTTTA TTGTTGAAAT	ACGCCGTGCC TGCGGCACGC	CTGTTCGCA	T TATCCGAAC A ATAGGCTTG	C ATCCGCTGTG G TAGGCGACAC	
2551	GTACACGCTG CATGTGCGAC	TGCGACCGC	r acggcctgt a tgccggaca	A TGTGGTGGA T ACACCACCT	T GAAGCCAATA A CTTCGGTTAT	
2601	AACTTTGGGT	GCCGTACCA	C GGTTACTTA	G CAGACTGGC	A TGATCCGCGC T ACTAGGCGCG	
	TGGCTACCGG	CGATGAGCG	A ACGCGTAAC	G CGAATGGTG	C AGCGCGATCG CG TCGCGCTAGC	
2701	ТАВТСВСССС	AGTGTGATC	A TCTGGTCGC T AGACCAGCC	T GGGGAATGA A CCCCTTAC	A TCAGGCCACG	
275	L GCGCTAATCA	CGACGCGCT	G TATCGCTGC	GA TCAAATCTO	GT CGATCCTTCC CA GCTAGGAAGG	
280	CGCCCGGTG(C AGTATGAAG G TCATACTTO	G CGGCGGAGGC GC GCCGCTC	CC GACACCAC GG CTGTGGTG	GG CCACCGATAT CC GGTGGCTATA	
						-

######################################							
2901 TGCCGRAATG GTCCATCAAA MANTGCTTI CGCTACCTGG AGAGAGCCGG AGGGCTTTAC CAGGTAGTT TTTACCCAAA GGATGGAC TCCTCCGCGG 2911 CCCCTGATCC TTTCCCAATA CGCCCACGGG ATGGGTAACA GTCTCCCGCG 3001 TTCCCTAAA TACTGCCAGG ATGGGTACAC GTCTTCCCGCG 3001 TTCCCTAAA TACTGCCAGG GTTCGTCA GTATCCCCGT TTACAGGGGG ANAGCGATT ATGACCGCC GCAAAGCAGT CATAGGCCA ANTGCCCCC 3001 GCTTCGTCTG GGACTGCGCT GATCAGTCGC TAGATAAATA TGATGAAAAAC CGAAGCACC CTGACCCCC CTGATCAGGC ATAGATTAATA TGATGAAAAAC CGAAGCACC CTGACCCCC CTGATCAGGC ATAGATTATA TGATGAAAAAC CGAAGCACC CTGACCCCC CTGATCAGGC ATAGATTATA CTACTTTTC 3101 GCGAACCCCCT GCACCCAC CTGATCAGGC ATAGATTATA TACATCTTTC 3102 CCGCACCCCT GCACCCAC CTGATCAGGC ATAGATTATA CTACTTTTC 3103 CCGCACCCCT GCACCCACAC AACCCACACACACACACACACACACACA	2851	TATTTGCCCG	ATGTACGCGC	GCGTGGATGA	AGACCAGCCC	TTCCCGGCTG	
2901 TGCCGRANTE GTCCATCARA ANATESCTIT CGCTACCTGE AGGACACGCC ACGCCTTAC CAGGACTTT TITLEGGANTA CGCCCACGGC ATGGGTAACA GTCTTGCCGC GGCGCTAGG AAAGGCTTAT GUGGGTGGGC TACCCATTGT CAGAACCGCC GGCGCTAGG AAAGGCTTAT GUGGGTGGGC TACCCATTGT CAGAACCGCC 3001 TTTCCCTAGA TACTGGCAGG CGTTTCGTCA GTATCCCCT TACAGGCGC AAAGCGATT ATGACCGCG GATCAGCTG CAGAGCGCA AATGTCCCCCT CGCAACCAGAC CCTGACCCAC CTAGTCAGCT CATATATAT ACTAGAAAC CGAACCAGAC CCTGACCCAC CTAGTCAGCT CAGATTATA ACTAGTATAC CGAACCAGAC CCTGACCCAC CTAGTCAGCT TACGCGCACCACC CCTTGGGCA CCAGCCCACAC CTAGTCACCAC AACCCGCTAT 3101 GGCAACCCGT GGTCGGCTTA CGGGGGTAT TTTGGGGATA CGCCGAACCAG CCCTTGGGCA CCAGCCCACAT TCGGCGCCATA AACCCCTAT GGGCTTCCT AGCCCTCAGA CAATCATCC CAGACCACACAA ACCCCCTAT CAGCCCTCAGA CAATCATCC CAGACCACACAA ACCCCCACAC CAGCCCCACAC CAGCCCCACA CAGCCACACAC CAGCCCCACAC CACCCCCACAC CTAGTCACCACCACACACACACACACACACACACACACAC					TCTGGTCGGG	AAGGGCCGAC	
ACGECTITAC CAGGIAGITI TITACGANA GCCAAGGG ATGGIACA TICTITGGGG GGGGACTIGG ANACCITAT GGGGGTCGC TACCCATTGT CAGGACCCC 3001 TITCGCTANA TACTGGCAGG CGTTTGGTCA GTATCCCGT TTACAGGGCG AAAGCGATTT ATGACCGTC GCAAAGCACT CATAGGGGCA AAATGTCCCGC 3051 GCTTCGTCTG GGACTGGGCTG GATCAGTCAC TATAGACGCCA 3051 GCTTCGTCTG GGACTGGCCTG CACAAGCACT CATAGACACCCC 3051 GCTCGTCTG GGACTGGCCTG CACAAGCACT CATAGACACCC CGAACCACCC GGCCCCAAA GCCGCTTA CACACCGCAACCAC CATAGTCACTTTTTG 3101 GGCAACCCCT GGTCGGCTTA CGGCGGTGAT TAGAGGACAC CCCTTGGCAA CACACCCAAA GCCGCGCATA AAACCGCTTA GCGCGAACCA CCCTTGGCAA CACACCCTAA GCCGCGCATA AAACCGCTTA GCGCGCACCAC 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCGC ACGCCCCATC 3251 CACCCCTTCA GGAACCACAA ACGCCACACACA ACGCCCCATC GGCCACCTCA GCACCCCAAAA ACGCCACACACA ACGCCCCATC GGCCACCTT GGAACCACACA ACGCCACACACA ACGCCACACACA ACGCCTCAAC GAACCACACA ACGCCACACACACACACACACACA					OCCUPA COMCC	7 C7 C7 CCCC	
2511 COCCTGATCC TITGCGATA CGCCCACGGG ATGGSTAACA STCTITGGGG GCGACTAGS AAAGCGTAT GGGGSTGGG TACCATTGT CAGAACGCC 3001 TITCGCTAAA TACTGGCAGG CGTTTGGTCA GGATCCCGT TACAGGGGG AAAGCGATT ATGACCCTC GCAAAGCAGT CATAGGGCG TACCATTGT CGAAGCAGT ATGACCCTC GCAAAGCAGT CATAGGGCGA AAAGCGCCA 3051 CCTTCGTCTC GGACCGGTG GATCAGCCC TAGTTAAATA TATGAGAAAC CGAAGCAGC CCTGACCCAC CTAGTCAGG ACTAATTTA ACTACTTTG 3101 GGCAACCCCT GGTCGCTA CGCGGGTGAT TTTGGCGCATA CGCGGAAGGA CCGTTGGGCA CCAGCCCAC CTAGTCAGG ACTAATTTA ACTACTTTG 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCGC ACGCGAACGA CCGTTGGGCA CCAGCCAACA ACTACTTGC CAGACCAGAA ACGCCTATA GCGCCGATC ACCCGTCAAC ACTACTTCC CAGACCAGA ACGCTGGAT GCGCGCATC ACCCGTCAAC ACTACTTCC CAGACCAGA ACGCCGAACGA ACGCCGCATC CAACCCCTAAC ACATACTTCC CAGACCAGAA ACGCCGAACGA ACGCCGCATC CAACCCCTAAC ACATACTTCC CAGACCAGAA ACGCCGAACGA ACGCCGAACGA CTAGACCACCAGAACA CACCAGCAGA ACGCTGAACGA CAGACCAGAC	2901	TGCCGAAATG	GTCCATCAAA	AAATGGCTTT	GCGATGGACC	TCTCTGCGCG	
GGGGCTAGG AAACGCTAT GGGGTGGGC TACCATTGT CAGAACGCC 3001 TTTCGCTAAA TACTGGCAGG GGTTTGGTCA GTATCCCGGT TTACAGGGGG AAAGCGATT ATGACCGTCC GCAAACACGAT CATAGAGGGCA AATGCCCGC 3051 GCTTCGTCTG GGACTGGGT GTTCGTCA GTATCAAATA TGATGAAAAC CGAAGCAGAC CCTGACCCAC CTAGTCAGCG ACTAATTAA TGATGATAAC CCAAGCAGAC CCTGACCCAC CTAGTCAGCG ACTAATTAATA TGATGATAAC 3101 GGCAACCCGT GGTCGGCTTA CGGCGGTGAT TTTGGCGATA CGCCGAACGA CCGTTGGCCA CCAGCCGAAT GCCCCCACTA AAACCGCTAT GCGGCTGGCT 3151 TCGCCGATTC TGTATGAACG GTCTGTCTT TGCCGACGC ACGCCGCATC ACGCCCTAGA CAAACCATCTCC CAGACCAGAA ACCCGCTAGCT ACGCCCTAGA CAAACCAGAAC CAGCCAGAA CACCAGGAC AGGCCGCATC ACGCCCACTAC ACAACCAGAA CACCAGGAC AGTTTTCCA GTTCCGTTTA GTCGCGACTG CCTCCGTTT GTGGTCGTCT TGCAAAAAGGT CAGGCCAAT 3251 TCCCGGGCAAA CCATCGAAGT GACCAGGAC AGTTTTCCA GTTCCGTTTA GTCGCCCCTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTACACGA AGCCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT ACGCCCCTTT GGTAGCTTCA CTGGTCGCCT ATGGACAAGG CAGTATCGCT TACAGGACTC CTGCACTGGA TGGTGCCCCC AAGGTAAACA GTTGATTGGA 3251 TACAGGACTC CTGCACTGGA TGGTGCCCCC AAGGTAAACA GTTGATTGGA CCCCCACTTCA CGGGGACCTA CAGCCACGCA CTACCATTC GCGCACCGTT 3251 GCGGCTAAGT GCCCTCGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CCCCCACTTCA CGGGGACCTA CAGCCAGCCG GGGCAACTCT GCCCACCTCA CCCCCTCAC CGGAGACCTA CAGCCAACC CAACCATCC CACCACTCCA CCCCGCACTCA CAGCCACCC GCACCGCAC GTCACAACCC CCCCGCACTCA CAGCCACCC GCACCGCAC GTCACAACCC CCCGGACCTC TGCCCCCC CACCCACCC CACCCCCCC CCCGGTAGT 3451 ACGCGTAGTC CACCCACCC GCACCGCAC GTCACAACCC GCCCGCTTACAC CCACCACTCCA CCCGGACCCT CTCCCCCC CCCCTTCCCCCC CCCCGTTCACC CCACCACTCCA 3501 CCCCCCACCC CACCCATCCC GCACCCCCTT TGGCGGAAAA ACCCACTCC CCCCGGCACCCC CCCGGACCCT CTCCCCCC CCTCTCCCCC CCCCTTTCACC CCCGACTCCCC CCCGGACCCT CTCCCCCC CCTTCCCCCC CCCCTTCACAC CCACCACCCCC CCCGGACCCT CTCCCCCC CCTTCACCC CACCACCCCC TCCCCCC CCCGGACCCT CTCCCCCC CCTTCACCC CACCACCCCC TCCCCCC CCCGGACCCT CTCCCCCC CCTTCACCC CACCACCCCC TCCCCCC CCCGGCACCCC CCCCCCCCC CCTCCACCC CCCCCCCCCC		ACGGCTTTAC	CAGGIAGIII	TITACCGAAA	GCOMIGGMEC		
GGGGCTAGG AAACGCTAT GGGGTGGGC TACCATTGT CAGAACGCC 3001 TTTCGCTAAA TACTGGCAGG GGTTTGGTCA GTATCCCGGT TTACAGGGGG AAAGCGATT ATGACCGTCC GCAAACACGAT CATAGAGGGCA AATGCCCGC 3051 GCTTCGTCTG GGACTGGGT GTTCGTCA GTATCAAATA TGATGAAAAC CGAAGCAGAC CCTGACCCAC CTAGTCAGCG ACTAATTAA TGATGATAAC CCAAGCAGAC CCTGACCCAC CTAGTCAGCG ACTAATTAATA TGATGATAAC 3101 GGCAACCCGT GGTCGGCTTA CGGCGGTGAT TTTGGCGATA CGCCGAACGA CCGTTGGCCA CCAGCCGAAT GCCCCCACTA AAACCGCTAT GCGGCTGGCT 3151 TCGCCGATTC TGTATGAACG GTCTGTCTT TGCCGACGC ACGCCGCATC ACGCCCTAGA CAAACCATCTCC CAGACCAGAA ACCCGCTAGCT ACGCCCTAGA CAAACCAGAAC CAGCCAGAA CACCAGGAC AGGCCGCATC ACGCCCACTAC ACAACCAGAA CACCAGGAC AGTTTTCCA GTTCCGTTTA GTCGCGACTG CCTCCGTTT GTGGTCGTCT TGCAAAAAGGT CAGGCCAAT 3251 TCCCGGGCAAA CCATCGAAGT GACCAGGAC AGTTTTCCA GTTCCGTTTA GTCGCCCCTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTACACGA AGCCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT ACGCCCCTTT GGTAGCTTCA CTGGTCGCCT ATGGACAAGG CAGTATCGCT TACAGGACTC CTGCACTGGA TGGTGCCCCC AAGGTAAACA GTTGATTGGA 3251 TACAGGACTC CTGCACTGGA TGGTGCCCCC AAGGTAAACA GTTGATTGGA CCCCCACTTCA CGGGGACCTA CAGCCACGCA CTACCATTC GCGCACCGTT 3251 GCGGCTAAGT GCCCTCGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CCCCCACTTCA CGGGGACCTA CAGCCAGCCG GGGCAACTCT GCCCACCTCA CCCCCTCAC CGGAGACCTA CAGCCAACC CAACCATCC CACCACTCCA CCCCGCACTCA CAGCCACCC GCACCGCAC GTCACAACCC CCCCGCACTCA CAGCCACCC GCACCGCAC GTCACAACCC CCCGGACCTC TGCCCCCC CACCCACCC CACCCCCCC CCCGGTAGT 3451 ACGCGTAGTC CACCCACCC GCACCGCAC GTCACAACCC GCCCGCTTACAC CCACCACTCCA CCCGGACCCT CTCCCCCC CCCCTTCCCCCC CCCCGTTCACC CCACCACTCCA 3501 CCCCCCACCC CACCCATCCC GCACCCCCTT TGGCGGAAAA ACCCACTCC CCCCGGCACCCC CCCGGACCCT CTCCCCCC CCTCTCCCCC CCCCTTTCACC CCCGACTCCCC CCCGGACCCT CTCCCCCC CCTTCCCCCC CCCCTTCACAC CCACCACCCCC CCCGGACCCT CTCCCCCC CCTTCACCC CACCACCCCC TCCCCCC CCCGGACCCT CTCCCCCC CCTTCACCC CACCACCCCC TCCCCCC CCCGGACCCT CTCCCCCC CCTTCACCC CACCACCCCC TCCCCCC CCCGGCACCCC CCCCCCCCC CCTCCACCC CCCCCCCCCC	2051	CCCCTCATCC	TTTCCCAATA	CCCCCACGCG	ATGGGTAACA	GTCTTGGCGG	
3001 TITCGCTARA TACTGCAGG CETTICGTCA GTATCCCCT THACAGGGCG AAAGCANTT ATGACCGTC GCAAAGCAGT CATAGGGCG AAAGCAGTC CGAAGCAGAC CCTGACCAC GTACAGCGC TATAGATATA TAGACCTTG CGAACCAGT GGTGGCTTA GGGGGTGAT TITGGGGATA GGGCGAACGA 3051 GGCAACCCGT GGTGGGCTTA GGGGGTGAT TITGGGGATA GGGCGAACGA 3101 GGCAACCCGT GGTGGGCTTA GGGGGTGAT TTGGGGATA GGGCGCAACGA CCGTTGGGCC CACACCACA GGCCCACTA AAACCCCTAT GCGGCTAGGA 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCGA ACGCGCGATC 42 AGCGCTCAG ACGACCAGAA ACGCCTGAG GCGCGCATC 43 AGCGCTCAG ACGACCAGAA ACGCCTGCG TGCGGCGTAG 43 AGCGCTCAG CCTTCGTTTT GTGGTGTGG TCAAAAAGGT CAAGCACAAC 44 AGCGCCTAG CCTTCGTTTT GTGGTGTGG TCAAAAAGGT CAAGAGCAAAT 45 TCCCGGCAAA CCATCCAGAGT GACCAGCAGA TACCTGTTCC GTCATAGCGA 45 TCCCGGCCAAA CCATCCAGAGT GACCAGCAGA TACCTGTTCC GTCATAGCGA 46 ATTCCCCAG GACCTGGAT GGTGGGGGT GGATGGTAAG CGCGGGCAA 47 TCCCCAGA GCACCTGGAT GTGGGCGCT GGATGGTAAG CGCGGGCAA 47 TCCCCAGA GACCTGGAT GTGGGCGCA CAACCATCC GGCGACCGTT 46 AGCGCCTTTG GGTAGCTTCA CACCGCGCA CAACCATCC GGCGCACAC 47 CCCCACTCA CGAACACCACCC GAGAGAGACA CATCCATTC GCCGACCGTT 47 CCCCACTCA CGAACACCACCC GGAGAGCACC GGCACACCATT GGCCACACGT 48 AGCGCTCTGAA GCCCTCTGGAT GTCGGTCACACTT GGCTCACAGT 49 AGCGGACTT GATGGGACCTA CACCGCACC GGGGAACTCT GGCTCACATT 49 AGCGCTAGAC CACCCAACC CGAGAGACCC GGGCACACCA TCCATTTCT GCACCACACCT 47 CCCCACCACCC GCACCACCC GGAGAGCCC CCCTTCACCACCC 47 CCCGCGACCCT CGTCCCCCA CACCCCATCCC GCCGCTTCACACCC 47 CCCGCGACCCT CGTCCCCCA GCCCCCTTTCACCC CCCGCGTTACT 47 CCCCACCACCC GCACTCCCC GCACCCCCTT TACCCCACCCC	2931	GGCGACTAGG	AAACGCTTAT	GCGGGTGCGC	TACCCATTGT	CAGAACCGCC	
AAAGCGATTT ATGACGATCG GAAAGCGT CATAGGGGCA AATGTCCCGC GCTTCGTCTE GGACTGGGTG GATCAGTCG TGATTAAATA TGATGAAAAC CCGAGCAGGAC CCTGACCCAC CTGGTCAGCG ACTAATTATA ACTACTTTTG 3101 GGCAACCCGT GGTGGGTTA CGGCGGTGAT TTTGGCGATA CGCCGAACGA CCGTTGGGCA CCAGCCGAAT CCGCCACTA TTTGGCGATA CGCCGGAACGA CCGTTGGGCA CCAGCCGAAT CCGCCACTA AAACCGCTAG CGGCGTAG 3151 TCGCCASTTC TGTATGAACG GTCTGGTCTT TCCCGACCGC ACGCCGCATC 3152 CAGCGCTGAC GCAACCAAA CACCACCAGA ACGCCTGGGT TCGGCGTAG 3153 CAGCGCTGAC GCAACCAAA CACCACCAGA ACGCCTGGGT TCCGGTTTA 3154 TCCGGGACAG CCATTCGTTT GTGGTCGTC TCAAAAAGGT CAAGGCAAAT 3155 TCCGGGCAAA CCATCGAAGT GACCACCAGA ACGCTGACTACAGCAAA 3161 TCCGGGGCAAA CCATCGAAGT GACCACCAGA TACCTGTTCC GTCATTACCGA 317 TAACGACCTC CTGCACTGAA TGGTGCGCT ATGCACACAGA ACCCCGCAAC 318 TAACGACCCC CTGCACTGA TGGTGCGCT ATGCACACAGC CAATTCGCT 317 TAACGACCTC CTGCACTGA TGGTGCGCT ATGCACACAGC CAATTCGCT 318 CGCCCCTTCA GGACCTCACCTC ACCACCCGCA CCTACCATTC GGCGACCGTT 319 CACCACCACCA CACCCCACCA AGGGAAACAC GTTGATTGAA 310 CTGCCTGAAC TACCGCACCC GACCACCAC AAGGGAAACAC GTTGATTGAA 310 CTGCCTGAAC TACCGCACCC GACCACCAC ACGCGGACCTC ACCACCTC 310 CACCACCAC GTTGGCTTCC CAACCGCACCA CACCACCAC 310 CACCACCAC GTTGGCTTCC CAACCGCACAA ACCTACCACC 310 CACCACCAC GTTGGCTTCC CAACCGCACAA ACCTACCACC 311 CACCACCAC GTTGGCTTCC CAACCGCACAAA ACCTACACTAC 311 CACCACCAC GTTGGCTTCC CAACCGCAAA ACCTACACTAC 312 CACCACCAC GTTGGCTTCC CAACCGCAAA ACCTACACTAC 313 CACCACCACC GACCACCC GACCACCAC CACCACCACAC 314 CACCACCACC GTTGGCTTCC CACCACCAC ACCACCACAA 315 CACCACCACC GTTGGCTTCCCCC GCACCACCAC ACCACCACAC 315 CACCACCAC GTTGGCTTCC CAACCGCAAA ACCTACACCACACC							
AAAGCGATTT ATGACGATCG GAAAGCGT CATAGGGGCA AATGTCCCGC GCTTCGTCTE GGACTGGGTG GATCAGTCG TGATTAAATA TGATGAAAAC CCGAGCAGGAC CCTGACCCAC CTGGTCAGCG ACTAATTATA ACTACTTTTG 3101 GGCAACCCGT GGTGGGTTA CGGCGGTGAT TTTGGCGATA CGCCGAACGA CCGTTGGGCA CCAGCCGAAT CCGCCACTA TTTGGCGATA CGCCGGAACGA CCGTTGGGCA CCAGCCGAAT CCGCCACTA AAACCGCTAG CGGCGTAG 3151 TCGCCASTTC TGTATGAACG GTCTGGTCTT TCCCGACCGC ACGCCGCATC 3152 CAGCGCTGAC GCAACCAAA CACCACCAGA ACGCCTGGGT TCGGCGTAG 3153 CAGCGCTGAC GCAACCAAA CACCACCAGA ACGCCTGGGT TCCGGTTTA 3154 TCCGGGACAG CCATTCGTTT GTGGTCGTC TCAAAAAGGT CAAGGCAAAT 3155 TCCGGGCAAA CCATCGAAGT GACCACCAGA ACGCTGACTACAGCAAA 3161 TCCGGGGCAAA CCATCGAAGT GACCACCAGA TACCTGTTCC GTCATTACCGA 317 TAACGACCTC CTGCACTGAA TGGTGCGCT ATGCACACAGA ACCCCGCAAC 318 TAACGACCCC CTGCACTGA TGGTGCGCT ATGCACACAGC CAATTCGCT 317 TAACGACCTC CTGCACTGA TGGTGCGCT ATGCACACAGC CAATTCGCT 318 CGCCCCTTCA GGACCTCACCTC ACCACCCGCA CCTACCATTC GGCGACCGTT 319 CACCACCACCA CACCCCACCA AGGGAAACAC GTTGATTGAA 310 CTGCCTGAAC TACCGCACCC GACCACCAC AAGGGAAACAC GTTGATTGAA 310 CTGCCTGAAC TACCGCACCC GACCACCAC ACGCGGACCTC ACCACCTC 310 CACCACCAC GTTGGCTTCC CAACCGCACCA CACCACCAC 310 CACCACCAC GTTGGCTTCC CAACCGCACAA ACCTACCACC 310 CACCACCAC GTTGGCTTCC CAACCGCACAA ACCTACCACC 311 CACCACCAC GTTGGCTTCC CAACCGCACAAA ACCTACACTAC 311 CACCACCAC GTTGGCTTCC CAACCGCAAA ACCTACACTAC 312 CACCACCAC GTTGGCTTCC CAACCGCAAA ACCTACACTAC 313 CACCACCACC GACCACCC GACCACCAC CACCACCACAC 314 CACCACCACC GTTGGCTTCC CACCACCAC ACCACCACAA 315 CACCACCACC GTTGGCTTCCCCC GCACCACCAC ACCACCACAC 315 CACCACCAC GTTGGCTTCC CAACCGCAAA ACCTACACCACACC	3001	ТТТСССТААА	TACTGGCAGG	CGTTTCGTCA	GTATCCCCGT	TTACAGGGCG	
3051 CITTCETCTE GGACTGGGTG GATCAGTCGC TGATTAAATA TGATGANAAC CGAAGCAGAC CCTGACCCAC CTAGTCAGCC ACCACATTTAT ACTACTTTE 3101 GGCAACCCGT GGTCGCCTTA CGGCGGTGAT TTTGGCGATA CCCGCAACCA CCGTTGGGCA CCACCCCATTA AAACCCGCTA CGCGCCATC CCGTTGGGCA CCACCCCATTA AAACCCGCTA CGCGCGTTGCT 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCGC ACCCCGCATC ACCGGCTCAGA ACCCGCACTA ACCGCCACTA AAACCCGCTA CGCCGCTAG 3151 TCGCCAGTTC GTATGAACA CACCACACA ACGCCTGCG TCCGGGTAG 3151 TCCGCGCACTG CCTTCGTTTT GTGGTCGT TCAAAAAAGT CACAGCAAAT 42 3251 TCCGGGCAAA CCATCGAACT GACCACGCGA TACCTGTCC TCAAAAAGT CACAGCAAAT 43 43 45 45 46 47 47 48 48 48 48 48 48 48 48	3001	AAAGCGATTT	ATGACCGTCC	GCAAAGCAGT	CATAGGGGCA	AATGTCCCGC	
CGARGCAGC CCTGACCCAC CRASTCAGCA ACTMATTAT ACTACTTTE 3101 GGCAACCCGT GGTCGGCTTA CGGCGGGAT TTTGGCGATA CGCCGAACGA CCGTTGGGCA CCAGCCGATT GCCGCCCACTA AAACCGCTAT GGGGCTTGCT 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCG ACGCCGCATC ACCGCTCAGC ACATACTTGC CAGACCAGAA ACGGCTGCG TCGGCCGTAG 3161 CAGCCCTGAC GGAAGCAAAA CACCAGCAGC AGTTTTTCCA GTTCCGTTTA GTCGGGACTG CCTTCGTTTT CTGGTCGTC TCAAAAAGGT CAAGCCAAAT 317 TCCGGGCAAA CCATCGAAGT GACCAGCAGA ATACTGTTCC GTCATAGCGA AGGCCCTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTACGCT 3251 TAACCAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA AGGCCCCTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CATCACTTC CTTCCTCAG GACCTGACA TCACCCGGA CCTACCATTC GGCGACCGTT 3251 GCGCTCAACT GCCTCGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA CGGAGCCTA CAGCCGGGA CCTACCATTC GGCGACCGTT CTTCCTCAG GACCTGACAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA TACGCAGCC GGAAGACCCC GGCAACTCT GGCACACTT CACCCACTCA CCGGAGCCTA CAGCCGAGC CCTTCACATTC GCCGACCGTT CTCCCCAGAC TACCGCAGCC GGAAGACCCC GGCAACTCT GGCACACTT TGCCCACTCA CTGGCTTCGC CCTCCCGGG CCCGTTGAGA CCGAGTGCTA 3451 ACGCTACTG ATGCCGTCCG CCTCTCCGGG CCCGTTGAGA CCGAGTGCTA 3501 GCGCCTGCA GCACCGAAC CACCCCTAC ACGCTCACACTT CCGCGCACCGT CTGCCTCCGC CCTCCCGTC CCGTCTAGTT 3501 GCGCCTGCA GCACTGCC GCACCCTTAT TGGAGTCACA CTGCGAGGGG 3511 CACCAGCTC ACGCCATCC GCATCTGACC ACCACCTTAA ACCTCACTTTT CGCGCCACCGT CATCCTCCAC GCACCCTTAA ACCTCACTTT ACCTAAAAAC 3601 CATCGAGCTG GCTAACACG GAACCCCTAA ATGCGCGTT CACCGCTAC 3611 TTCACAGAT GTCGATTGGC GTTAGACCACTTTA ATTCGCGCT CACCGCTTAC 3701 GATCAGTTCA CCCGTCACC GCATCATTTTT TTTTTTTTTT							
CGARGCAGC CCTGACCCAC CRASTCAGCA ACTMATTAT ACTACTTTE 3101 GGCAACCCGT GGTCGGCTTA CGGCGGGAT TTTGGCGATA CGCCGAACGA CCGTTGGGCA CCAGCCGATT GCCGCCCACTA AAACCGCTAT GGGGCTTGCT 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCG ACGCCGCATC ACCGCTCAGC ACATACTTGC CAGACCAGAA ACGGCTGCG TCGGCCGTAG 3161 CAGCCCTGAC GGAAGCAAAA CACCAGCAGC AGTTTTTCCA GTTCCGTTTA GTCGGGACTG CCTTCGTTTT CTGGTCGTC TCAAAAAGGT CAAGCCAAAT 317 TCCGGGCAAA CCATCGAAGT GACCAGCAGA ATACTGTTCC GTCATAGCGA AGGCCCTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTACGCT 3251 TAACCAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA AGGCCCCTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CATCACTTC CTTCCTCAG GACCTGACA TCACCCGGA CCTACCATTC GGCGACCGTT 3251 GCGCTCAACT GCCTCGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA CGGAGCCTA CAGCCGGGA CCTACCATTC GGCGACCGTT CTTCCTCAG GACCTGACAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA TACGCAGCC GGAAGACCCC GGCAACTCT GGCACACTT CACCCACTCA CCGGAGCCTA CAGCCGAGC CCTTCACATTC GCCGACCGTT CTCCCCAGAC TACCGCAGCC GGAAGACCCC GGCAACTCT GGCACACTT TGCCCACTCA CTGGCTTCGC CCTCCCGGG CCCGTTGAGA CCGAGTGCTA 3451 ACGCTACTG ATGCCGTCCG CCTCTCCGGG CCCGTTGAGA CCGAGTGCTA 3501 GCGCCTGCA GCACCGAAC CACCCCTAC ACGCTCACACTT CCGCGCACCGT CTGCCTCCGC CCTCCCGTC CCGTCTAGTT 3501 GCGCCTGCA GCACTGCC GCACCCTTAT TGGAGTCACA CTGCGAGGGG 3511 CACCAGCTC ACGCCATCC GCATCTGACC ACCACCTTAA ACCTCACTTTT CGCGCCACCGT CATCCTCCAC GCACCCTTAA ACCTCACTTT ACCTAAAAAC 3601 CATCGAGCTG GCTAACACG GAACCCCTAA ATGCGCGTT CACCGCTAC 3611 TTCACAGAT GTCGATTGGC GTTAGACCACTTTA ATTCGCGCT CACCGCTTAC 3701 GATCAGTTCA CCCGTCACC GCATCATTTTT TTTTTTTTTT	3051	GCTTCGTCTG	GGACTGGGTG	GATCAGTCGC	TGATTAAATA	TGATGAAAAC	
CCGTTGGCCA CCAGCCGAAT GCCCCCCATA ANACCGCTAT GCGGCTTGCT 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCGC ACGCCGCATC AGCGGTCAAG ACATACTTGC CAGACCAGAA ACGCTGGCG TGCGGCGTAG 3251 CAGCGTCAAC GGAAGCAAAA CACCAGCAGC AGTTTTTCCA GTTCCGTTTA 11 GTGCCGCACTG CCTTGGTTTT TGGGTCCTGC TCAAAAAGGT CAGGCCAAAT 12 TCCGGGCAAA CCATCGAAGT GACCAGCGAA TACCTGTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 3251 TCCGGGCAAG CCATCGAA TGGTGGCGCT GGATGGTAAG CCGCTGCGAA ATTGCTCGAG GACGTAACCT ACCACCGCGA CCTTACCATTC GGCGACCGTT 3351 TAACGAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT 3451 GCGCTGAAC TACCGCAGCC GGAGAGCGCC GGCGAACTCT GGCTCACACTT 3451 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGCGAACTCT GGCTCACAGT 3451 ACGCCTGAC GAACCGAAGC CCGCACGGG CCCCTTGACA CCGAGTGTCA 3451 ACGCCTAGC CAACCGAAGC CCACCCCATG GTCAGAAGCC GGGCACTCA TGCGCCATCAC GTTGGCTTGC GCTGGCGGAA ACCTGACTGT 3501 GCGCTGGCA GCAGTGGCCT CTGGCGGAAA ACCTCAGTGT 3501 GCGCTGCA GCAGTGGCCT CTGGCGGAAA ACCTCAGTGT 3501 GCGCCTGCC ACGCCAACC CAACCGCAAT GTCAGTACCC CCGCGACCTC GGTCACCGCA GACCCCCTTT TGGATCACA CCGAGTGTCA 3501 GCGCCTGCC ACGCCATCC GCATCGACC ACCACGGAA TGGATTTTTC CGCCGCACCTC CGTCACCGCA GACCCCCTTT TGGATCACA CCGAGTGTCA 3501 CATCCAACCG GTAGACCG CAACCGCATT TGGGTCACAC CCACCGTATTTC CTGCCCACCC CACCCATCC CAACCGCAAT TACCCCACAC TCACCCCC CCGCGCACCG GCGGAGGG CCTAGACCG CAACCGCAAT TACCTCAAAAAAC 3601 CATCCAACCG GTAATAACC GTTGGCATT TACCCCCAC TCACCCCCC CTACCCCAC CCATTATTCC CAACCGTTAA ATTGCCGCTC ACCCCCCCC CAACACTCA CCCCATTATTCT CAACCGCCAC TCACCCCACCC		CGAAGCAGAC	CCTGACCCAC	CTAGTCAGCG	ACTAATTTAT	ACTACTTTTG	
CCGTTGGCCA CCAGCCGAAT GCCCCCCATA ANACCGCTAT GCGGCTTGCT 3151 TCGCCAGTTC TGTATGAACG GTCTGGTCTT TGCCGACCGC ACGCCGCATC AGCGGTCAAG ACATACTTGC CAGACCAGAA ACGCTGGCG TGCGGCGTAG 3251 CAGCGTCAAC GGAAGCAAAA CACCAGCAGC AGTTTTTCCA GTTCCGTTTA 11 GTGCCGCACTG CCTTGGTTTT TGGGTCCTGC TCAAAAAGGT CAGGCCAAAT 12 TCCGGGCAAA CCATCGAAGT GACCAGCGAA TACCTGTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 3251 TCCGGGCAAG CCATCGAA TGGTGGCGCT GGATGGTAAG CCGCTGCGAA ATTGCTCGAG GACGTAACCT ACCACCGCGA CCTTACCATTC GGCGACCGTT 3351 TAACGAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT 3451 GCGCTGAAC TACCGCAGCC GGAGAGCGCC GGCGAACTCT GGCTCACACTT 3451 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGCGAACTCT GGCTCACAGT 3451 ACGCCTGAC GAACCGAAGC CCGCACGGG CCCCTTGACA CCGAGTGTCA 3451 ACGCCTAGC CAACCGAAGC CCACCCCATG GTCAGAAGCC GGGCACTCA TGCGCCATCAC GTTGGCTTGC GCTGGCGGAA ACCTGACTGT 3501 GCGCTGGCA GCAGTGGCCT CTGGCGGAAA ACCTCAGTGT 3501 GCGCTGCA GCAGTGGCCT CTGGCGGAAA ACCTCAGTGT 3501 GCGCCTGCC ACGCCAACC CAACCGCAAT GTCAGTACCC CCGCGACCTC GGTCACCGCA GACCCCCTTT TGGATCACA CCGAGTGTCA 3501 GCGCCTGCC ACGCCATCC GCATCGACC ACCACGGAA TGGATTTTTC CGCCGCACCTC CGTCACCGCA GACCCCCTTT TGGATCACA CCGAGTGTCA 3501 CATCCAACCG GTAGACCG CAACCGCATT TGGGTCACAC CCACCGTATTTC CTGCCCACCC CACCCATCC CAACCGCAAT TACCCCACAC TCACCCCC CCGCGCACCG GCGGAGGG CCTAGACCG CAACCGCAAT TACCTCAAAAAAC 3601 CATCCAACCG GTAATAACC GTTGGCATT TACCCCCAC TCACCCCCC CTACCCCAC CCATTATTCC CAACCGTTAA ATTGCCGCTC ACCCCCCCC CAACACTCA CCCCATTATTCT CAACCGCCAC TCACCCCACCC							
151 TCGCCAGTTC TGTATGAACS GTCTGGTCTT TGCCGACCGC ACGCCGCATC ACGCGTCAAG ACATACTTGC CAGACCAGAA ACGGCTGCGC TGCGGCGTAG 101 CACCGCTCAC GGAAGCAAAA CACCAGCAGC AGTTTTCCA GTTCCGTTTA GTGCCGACTG CCTTCGTTTT GTGGTCGTGG TCAAAAAGGT CAAGGCAAAT 102 GTCGCGACTG CCTTCGTTTT GTGGTCGTGG TCAAAAAGGT CAAGGCAAAT 103 TCCGGGCAAA CCATCGAAGT GACCAGCGAA TACCTGTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTTCA CTGGTCGCCTT ATGACAAAGG CAGTATCCCT 103 TAACGAGCTC CTGCACTGGA TGGTGGGCGT GGATGGGTAAGC CCGCTGCCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT 104 CTGCCTGAAC TACCGCAGCC ACGCCGACCTCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA CGGGGGCCTA CACCACCGCA CCTACCATTC TCAACTAACTT 105 CTGCCTGAAC TACCGCAGCC GGAGAGCCC CGGCAACTCT GGCTCACAGT CGCCACTCAC CGGGGACCTA CAGCCAGGTG CCCGTTGAGA CCGAGTGTCA 1361 ACGGTTAGTG CAACCGAACG CCACCGCATG GTCAGAAGCC CGGCCACTCA TGCCCATCAC GTTGGCTTGC GCTGGCGGTA CAGTCTTCGG CCCGTGTAGT 1351 GCGCCTGCA GCAGTGGCC CGCCGCACCGCAT GAGCACCCCAC CGCGGACCCT CGTCACCCCA GACCGCATT TGGAGCACC CCGCGGGGGGGG 1351 GCCCCCTCC ACGCCATCC CCATCTGACC ACCACCGAAA TGGATTTTG CGCGCATCAC GTGCGCTACC GCTACTCACC CACCCCAAA TGGATTTTTG CGCGCACGG TGCGGTAGGC CGTAGACTG TGGGCGTTT ACCTAAAAAAC 1361 CATCGAGCTG GGTAATAACC GTTGGCAATT TACCACCAGA TGGATTTTTG CGCGCCAGGC GGCGAACCCC CATTATTTTG CAACCGAAA TGGATTTTTG CGCGCCAGGC GGCGAGGCGC CGTAGACTG TGGGCGTTT ACCTAAAAAAC 1361 TTTCACAGAT GTGGATTGGC GATAAAAAAC ACCACCGAAA TGGATTTTTG CTACCTCGAC CCATTATTTG CAACCGTATA ATTGCCGCT AGGCCTTC CTACCCAAGT GGCACCGTTAA ATTGCCGCT TACCTAAAAAAC 1361 GATCAGTTCA CCCCTGCACC GCTGGATAAC GACCTTCAC GCGCGCGCGC 13701 GATCAGTTCA CCCCTGCACC GCTGGATAAC GACCTTCAC GCGCGCGCC 13701 GATCAGTTCA CCCCTGCACC GCTGGATAACC GACCTTATA CCTCAAACACC CTAGTCCAAC GGCCACCGTACC CCTGGATAACC GACCTTATA TTGCCGCCCCC CTGGCCCTAACC GCCCTTATC CTCTAACCGC TAATCACTTCC CTGGCCCTAA CCCCTAACCG CACCGTATC CCCGCGCGCCC 13701 GACCACCTTA CCCCTGCACC GCTGGATAACC GACCTTATA CCCCAACCGCAA 13701 GACCACCTTCACC CCTGGCACC GCTGGATACC GCCCGCGCGCC 13701 GACCACCTTACC CCCTGCACC GCTGGATACC CCTGAACCCTTC GCCCCCCCCCC	3101	GGCAACCCGT	GGTCGGCTTA	CGGCGGTGAT	TTTGGCGATA	CGCCGAACGA	
AGGGGTGAG ACATACTTGC CAGACCAGAA ACGGTGGCG TGGGGGTAG 1301 CAGCGCTGAC GGAAGCAAAA CACCAGCAGC AGTTTTTCCA GTTCCGTTTA 147 GTCGCGACTG CCTTCCTTTT GTGGTCGTCG TCAAAAAGGT CAAGGCAAAT 147 ACGAGCTGAAC CATCGAACT GACCAGCAAA TACCTGTTCC GTCATAGCGA 1501 AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 151 ATCCTGGAG GACGTGACT ACCACCGCGA GATGGTAAG CCGCTGGCAA 151 ACGAGCTC CTGCACTGCA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA 151 CCGCATCTA CGGAGACCTA CACCCGCGAA CTTCCATTTC GGGACCCGTT 152 GCGCTGAACT TACCGCAGCC GAGAGGCCC GGCAACTCT 153 ACGCGTAGTG CAACCGAACC CGACGAGGCC GGCAACTCT 154 CACCACACC GTTGGCTTC CCTCCACC AAGGTAAACA CTTCATTGAA 155 ACGCGTAGTC CAACCGAACC CGACCACCAC CCCCCTTGACA CCGAGTGTCA 156 ACGCGTAGTC CAACCGAACC CGACCGCATC CACCCCCACCACCACCACCACCACCACCACCACCACCA		CCGTTGGGCA	CCAGCCGAAT	GCCGCCACTA	AAACCGCTAT	GCGGCTTGCT	
AGGGGTGAG ACATACTTGC CAGACCAGAA ACGGTGGCG TGGGGGTAG 1301 CAGCGCTGAC GGAAGCAAAA CACCAGCAGC AGTTTTTCCA GTTCCGTTTA 147 GTCGCGACTG CCTTCCTTTT GTGGTCGTCG TCAAAAAGGT CAAGGCAAAT 147 ACGAGCTGAAC CATCGAACT GACCAGCAAA TACCTGTTCC GTCATAGCGA 1501 AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 151 ATCCTGGAG GACGTGACT ACCACCGCGA GATGGTAAG CCGCTGGCAA 151 ACGAGCTC CTGCACTGCA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA 151 CCGCATCTA CGGAGACCTA CACCCGCGAA CTTCCATTTC GGGACCCGTT 152 GCGCTGAACT TACCGCAGCC GAGAGGCCC GGCAACTCT 153 ACGCGTAGTG CAACCGAACC CGACGAGGCC GGCAACTCT 154 CACCACACC GTTGGCTTC CCTCCACC AAGGTAAACA CTTCATTGAA 155 ACGCGTAGTC CAACCGAACC CGACCACCAC CCCCCTTGACA CCGAGTGTCA 156 ACGCGTAGTC CAACCGAACC CGACCGCATC CACCCCCACCACCACCACCACCACCACCACCACCACCA							
### ACCCCCCCC GGAAGCAAAA CACCAGCAGC AGTTTTTCA GTTCCGTTTA	Contract Contract	TCGCCAGTTC	TGTATGAACG	GTCTGGTCTT	TGCCGACCGC	TCCCCCCTAG	
3201 CAGGGCTGAC GGAAGCAAAA CACCAGCAGC AGTTTTTCA GTTCCTTTA GTCGCGACTG CCTTCGTTTT GTGGTCGTC TCAAAAAGGT CAAGGCAAAT 3251 TCCGGGCAAA CCATCGAAGT GACCAGCGGA TACCTCTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 3301 TAACGAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT 3310 CAGCGTGAAGT GCCTCTGGAT GTGGCTCCAC AAGGTAAACA GTTGATTGAA LL GCCCACTTCA CGGAGCCTA CACCACGGGG TCCATTCT GCCCACCATT AGCGACTTC GAGCGACCTA CACCAGGGGG TCCATTCTTC CAACTAACCTT 3401 CTGCCTGAAC TACCGCAGCC GGAGGACGCC GGGCAACTCT GGCTCACAGT CACGGACTTG ATGGGGTCGG CCTCTCGCGG CCCGTTGACA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TCCGCATCAC GTTGGCTTCC GCTGCGGTAC CAGTCTTCGG CCCGTTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCC CGCGGACCGT CGTCACCGCA GACCGCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCCGGTCC ACGCCATCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCCTGAC GCGGTAGGC GTTGGCCAAT TAACCGCAAA 3661 CATCCAAGGT GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTT CCGGCCATGAC CCATTATTCG CAACCGTAAA ATTGGCGTT ACCTAAAAAC 3661 TTTCACAGAT GTGGATTGAC GAACCGTAAA ATTGCGCAC ACTCCGAAAG 3701 GATCAGTCAA CTGCGACC GCTGGATAAAAAAC AACTCCTGAC GCCGCGCCCG 3701 GATCAGTCAA CCCGTGACC GCTGGATAAAAAAC AACTCCTGAC GCCGCGCCCC CTAGTCAACT GCCCGCACC GCTGGATTAT TTTTTTTTTT		AGCGGTCAAG	ACATACTTGC	CAGACCAGAA	ACGGC1GGCG	IGCGGCGIAG	
GCGCGGACA CCATCGATT GTGCCGCGA TACCTGTTC GTCATAGCA AGGCCCGTTT GGTAGCTCA CTGGTCGCTT ATGGACAAG CAGTATCGCT TACCGGCAAA CCATCGAAGT GACCAGCGAA TACCTGTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT TACCGCGCGAC GACGGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT CCGCGGGAAGT GCCTCTGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTCT CAACTAACTT CGCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTCT CAACTAACTT CGCCACTTCA CGGAGACCTA CAGCGAGGTG CCCTTGAGA CCGAGTGTCA ATGGCGTAGAC TACCGCAGC GGAGAGCGC GGGCAACTCT GGCTCACAGT GACGGACTTG ATGGCGTCGG CCCCTTGAGA CCGAGTGTCA ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCAACTCA TCGCGATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT CGCCCTGCAC GTTGGCTTCC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT CGCCCGCACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCCGCGCACGG TGCCGTTAGCC ACCACCGCAA TACCTAAAAAC CATCGAGCTG TGCGCTTAGC GACCGCATT TAACCGCCAA TACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAA TACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAA TACCTAAAAAC 3601 CATCGAGCTG CCATTATTCG CAACCGTAAA ATTGGCGGTC AGCCCTTTC GTAGCTCGAC CCATTATTCG CAACCGTAAA AATTGCCGCA CGCGCGGCC AAAGTGTCTA CACCTAACCG CATTATTTTG TTGACGACTG CGCGCCACGCG 3701 GATCAGTTCA CCCCTGCACC GCTGGATTAA CACTATGCGC TACGCTTTC CTGGCGCTAA CTCGGGATTGC CGTGGATGAC GCCGCGGCC 3701 GATCAGTTCA CCCCTGCACC GCTGGGTCAA CACCTTCCCCCCCCCC					a Common CCa	CTTCCCTTTT	
3251 TCCGGGCAAA CCATCGAAGT GACCAGCGAA TACCTGTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 3301 TAACGAGCTC CTGCACTGGA TGGTGGCGT GGATGGTAAG CCGCTGGCAA ATTGCTCGAG GACGTCACCT ACCACCGGA CCTACCATTC GGCCACCGTT 3251 GCGGTGAAGT GCCTCTGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA GCCCACTTCA CGGAGACCTA CAGCGAGGT TCCATTTGT CAACTAACTT 3451 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GGCTCACAGT GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCTAGTG CAACCGAACG CCTCCCGCG CCCGTTGAGA CCGAGTGTCA 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCACCCC CGCGGACCGT CGTCACCGCA GACCGCTTT TGGAGTCACA CTGCCAGGGG 351 GCCCCTGCCA GCCCATCCC GCATCGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TCCGTAAGAG CTGACCGTT TACCGCAGG CAGCGCTTC ACCTGAAAAAC 3601 CATCGAGCT GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGCGTTCC GTAGCTCGAC CCATTATTCG CAACCGTAAA ATTGGCGGT AGTCGAAAG 361 TTTCACAGAT GTGGATTGC GTTAGACTA ATTGGCGGT AGTCGAAAG 361 TTTCACAGAT GTGGATTGC CATTTTTTT TTGACGACTG CAGCGCACCG 3701 GATCAGTCA CCCTGACCG CTATTTTTT TTGACGACTG CAGCGGACCGC 3701 GATCAGTCA CCCGTGACC CCTGGTGAA CGCTGGAAA CCCTCAGCCAACCCTTAGTCAACCG CTATTTTTT TTGACGACTG TAAGTGAACC CTAGTCAAGT GGGCACTG CCCTGGTCAC CCTGGTCGCA ACCCTGAACCG ATTCACTTCG 3751 GACCGCCATT GACCCTAACC CCTGGTCAA CGCTGGAAA CCCGCGGGCC CTAGTCAAGT GGGACCTGAC CCTGGGTCAA ACCGTGGAAA CCCGCGGGCC CTGGGCCGTAA CTGGGATTGC CCTGGGTCAA ACCGTGGAAA CCCGCGGCCC CTGGGCCGTAA CTGGGATTGC CCTGGGTCAA ACCGTGGAAA CCCGCGCCCCGGCC CTGGGCCGTAA CTGGGATTGC CCTGGGTCAA ACCGTGGAAA CCCGCGCCCGGCC CTGGGCCGTAA CTGGGATTGC CCTGGGTCAA ACCGTGGAAA CCCGCGCCCCGGCC CTGGGCCGTAA CTGGGATTGC CCTGGGCCAA CCCCTGGGCCCCCCCCCGGCCCCCCCCCC		CAGCGCTGAC	GGAAGCAAAA	CACCAGCAGC	TCAAAAAGGT	CAAGGCAAAT	
13251 TCCGGGCAAA CCATCGAAGT GACCAGCGAA TACCTGTTCC GTCATAGCGA AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT 13301 TAACGAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT 1351 GCGGTGAAGT GCCTCTGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA CGGAGCCTA CAGCGAGGTG TTCCATTGT CAACTAACTT 1351 CGCCCACTTCA CGGAGCCC CAGCGAGGTG TTCCATTGT CAACTAACTT 1351 ACGCCTGAAC TACCGCAGCC GAGGAGCGC GGCGAACTCA CTGCCTGAAC ATGGCCTCGC CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 1351 ACGCCTAGTG CAACCGAACG CGACGGCATCA CAGCTAAGC CCGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGGAAA CCTCACTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 1351 GCGCCTGGCA GCAGTGGCG TCTGCGCGAAAA CCTCACTGT GACGCTCCCC CGCGCACGG TGCGGTAAGAG CGATCGACCGCAAA TGGATTTTTG CGGCCCTGGC ACGCCATCC GCACCGCAAA TGGATTTTTG CGGCCTAGC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGCCTCACAAG 1601 CATCGAGCT GGTAATAAGC GTTGGCCAATT TAACCGCCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGT AGCTCGCAAAG 1701 GATCAGATCA CCCGTGCACC CTATTTTTT TTGACGACTG CGCGGCACCGC 1701 GATCAGTCA CCCGTCACC CTATTTTTTT TTGACGACTG CAACCGTGCCAAAC 1701 GATCAGTCA CCCGTGCACC CTATTTTTTT TTGACGACTG CAACCGTGCCCAAACCTTAACACC CTAGTCAAACC CTAGTCAAACC CTAGTCAAACC CTAGTCAAACC CTAGTCAACCC CTAGTCAACCCACCTAACCC CTAGTCAACCCACCTAACCCCCACCTAACCC CTAGTCAACCCCACCTAACCCCCACCTAACCCCACCACCTAACCC		GICGCGACIG	CCIICGIIII				
AGGCCCGTTT GGTAGCTTCA CTGGTCGCTT ATGGACAAGG CAGTATCGCT TAACGAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT GCGGTGAAGT GCCTCTGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA CGCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT AG101 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GCCTCACAGT GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGATG GTCAGAAGCC GGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCCGGTCC ACGCCATCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTACACACTG TCGCACAGG 3601 CATCGAGCT GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTACCTCGAC CCATTATTCC CAACCGTTAA ATTGGCGGTC ACTCCGAAAG 361 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCC AAAGTGCTAA CACCTAACCG CTATTTTTTTTTTTTTTT		macacacan n		CACCACCGAA	ሞልርርጥርጥጥርር	GTCATAGCGA	
TARCGAGCTC CTGCACTGGA TGGTGGCGCT GGATGGTAAG CCGCTGGCAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCCACCGTT GCGCACTTCA CGGAGACCT ACCACCGCGA CCTACCATTC GGCCACCGTT GCGCACTTCA CGGAGACCTA CACGAGGTG TTCCATTTGT CAACTAACTT GCGCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT ACGCTGAACT GACCGAACG CGGAGAGCGCC GGCAACTCT GGCTCACAGT GACGGACTTG ATGGCGTCGG CCTCTCCGCG CCCGTTGAGA CCGACTGTCA ACGCTTAGTC ATGGCGTCGC CTGCCGCACC CAGCCACTCA CAGCCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGCTCTCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCCAGGGG 3551 GCCCCGTCCC ACGCCATCCC GCACCTGACCA ACCACGCGAAA TGGATTTTTG CGCGCACGG TGCGGTAGGG CGTAGACTGG TGGTCGCTACCGC TCGCGAGGGG 3651 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCCAAAG 3651 TTTCACACAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCGCCGC	:	TCCGGGCAAA	CCATCGAAGI	CTGGTCGCTT	ATGGACAAGG	CAGTATCGCT	
TAAGGACTC CTGCACTGGA TGGTGGCGCT GATGGTAGA CCGCTGGAAA ATTGCTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT THE TIGGTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT GCGCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT THE TIGGTCGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GGCTCACACT GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TGCCCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTG TGGTCGTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTAAA ATTGGCGTC AGTCCGAAAG 361 TTTCACAGAT GTGGATTGC GATAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTT TTGACGACTG CGGCGACGCG 3701 GATCAGTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGCC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GACCCTAACCG CTTGTTTTTTT GTTGACGACTG CGGCGCGCC 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCCAAGT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGCGCGCC 3701 GATCAGTTCA CCCGTGCACC CCTGGGTCGA ACGCTGGAAG GCGCGCGCC CTGGGCGTAA CTGGGATTGC GCACCTATTC CTGTAACCGC ATTCACTTCC 3710 GACCCCCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGCCGCCCCGC 3710 GACCCCCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGCCCCCGGC CTGGGCGTAA CTGGGATTGC GCACCTATTC CTGTAACCGC ATTCACTTCC CTGGGCGTAA CTGGGATTGC GCACCTATTC CTGTAACCGC ATTCACTTC CGCCCCCGG		AGGCCCGIII					
ATTECTCGAG GACGTGACCT ACCACCGCGA CCTACCATTC GGCGACCGTT GCGGTGAAGT GCCTCTGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA GCCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT AA01 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GGCTCACAGT GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGACTGTCA ACGCGTACTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGGAA ACCTCACTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCATT TGGAGTCACA CTGCGAGGGG 3501 GCGCCTGCCA GCACTGCCC GCACCGCTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTG TGGTCGCTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTT TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCAC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAACT GGGCACTGC CGCTGGATACC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTTGTTTTTT GTGACGACTG CGGCGCACGCG 3701 GATCAGTTCA CCCGTGCAC CCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAACT GACCCTAACG CCTGGATAAC GACCATTTG CTGTAACCGC ATTCACTTCC 3751 GACCGCCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCCCCGG CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCCCCGG		ሞ አ አ ርር አ ርርጥር	CTGCACTGGA	TGGTGGCGCT	' GGATGGTAAG	CCGCTGGCAA	
GCGGTGAAGT GCCTCTGGAT GTCGCTCCAC AAGGTAAACA GTTGATTGAA GCCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT GCCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT GCCCACTTCA CGGAGACCTC GAGAGACGCC GGGCAACTCT GGCTCACAGT GACCGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCACAAGCC GGCACACTCA TGCGCATCAC GTTGGCTTCC GCTGGCGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCATT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGGTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGGC AAAGGTCTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGCCCATACC GCACCTATTTT CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGACG ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCCGCGGGCC CTGGGCCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCCGCGCGCC 3761 GACCCCCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCCGCGGGCC CTGGGCCATT GACCCTAACG CCTGGGTCGA ACGCTGAAG GCCGCGCGCC 3761 GACCCCCATT GACCCTAACG CTGGATAAC GACATTGCC TAAGTGAAGC CTGGGCCATT GACCCTAACG CCTGGGTCGA ACGCTGAACG CCGCCCCGGCCCCGGCCCCGCCCC	8	ATTGCTCGAG	GACGTGACCT	ACCACCGCGA	CCTACCATTC	GGCGACCGTT	
CGCCACTTCA CGGAGACCTA CAGCGAGGTG TTCCATTTGT CAACTAACTT 3401 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GGCTCACAGT GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGGAAA ACCTCAGTGT GACGCTCCC CGCGGACCGT CGTCACCGCA GACCGCATA TGGAGTCACA CTGCGAGGGG 3501 GCGCCTGCCA GCAGCACCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCACGG TGCGCTAGCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGC GTTGGCAATT TAACCGCCAG TCAGAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTTTTTTTTTT	<u> </u>		- 				
2401 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GGCTCACAGT GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCATT TGGAGTCACA CTGCGAGGGG 3501 GCGCCTGCCA GCACTCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTACCTCCAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACCCC 3701 GATCAGTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAACT GGCCACTGG CGACCTATT CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAG GCGGCGGCC CTGGGGGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAG GCGGCGCC CTGGGGGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	3351	GCGGTGAAGT	GCCTCTGGAT	GTCGCTCCAC	AAGGTAAACA	GTTGATTGAA	
3401 CTGCCTGAAC TACCGCAGCC GGAGAGCGCC GGGCAACTCT GGCTCACAGT GACGGACTTG ATGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCCCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGCC GAACACGTTAA ACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTT TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTG CGACCTATTT CTGACACCG ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGCC CTGGGCGTAAA CTGGGGATTGC GGACCCAGCT TGCGACCTTC CGCCCCCGG		CGCCACTTCA	CGGAGACCTA	CAGCGAGGT	TTCCATTTGT		
GACGGACTTG ATGGCGTCGG CCTCTCGCGG CCCGTTGAGA CCGAGTGTCA 3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCCAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTT TTGACGACTG CGGCGACCGG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GACCCTAACG CCTGGGTCGA ACCCTGGAAG GCGGCGGCC 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACCCTGGAAG GCGGCGGCC CTGGGCGTAA CTGGGGATTGC CGGACCCAGCT TGCGACCTTC CGCCCCGG							
3451 ACGCGTAGTG CAACCGAACG CGACCGCATG GTCAGAAGCC GGGCACATCA TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCCGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTT TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTT CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGCGGGGCC CTGGGCGTAA CTGGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG		CTGCCTGAAC	TACCGCAGCC	GGAGAGCGC	C GGGCAACTC1	GGCTCACAGT	
TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCCGTAA CTGGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	2 2 22	GACGGACTTG	ATGGCGTCGG	G CCTCTCGCGC	CCCGTTGAGE	CCGAGIGICA	
TGCGCATCAC GTTGGCTTGC GCTGGCGTAC CAGTCTTCGG CCCGTGTAGT 3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCCGTAA CTGGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG						CCCCACATCA	
3501 GCGCCTGGCA GCAGTGGCGT CTGGCGGAAA ACCTCAGTGT GACGCTCCCC CGCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGCGGGCC CTGGGCGTAA CTGGGATTGCC GGACCCAGCT TGCGACCTTC CGCCCCCGG	3451	ACGCGTAGTG	CAACCGAACC	CGACCGCATO	CAGTCTTCG	G CCCGTGTAGT	
GCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG		IGCGCATCAC					
GCGGACCGT CGTCACCGCA GACCGCCTTT TGGAGTCACA CTGCGAGGGG 3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	2501	CCCCCTCCCT	CCACTGGCG	r CTGGCGGAA	A ACCTCAGTG	r GACGCTCCCC	
3551 GCCGCGTCCC ACGCCATCCC GCATCTGACC ACCAGCGAAA TGGATTTTTG CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCCGG	3301	CGCGGACCG	CGTCACCGC	A GACCGCCTT	T TGGAGTCAC	A CTGCGAGGGG	
CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG							
CGGCGCAGGG TGCGGTAGGG CGTAGACTGG TGGTCGCTTT ACCTAAAAAC 3601 CATCGAGCTG GGTAATAAGC GTTGGCAATT TAACCGCCAG TCAGGCTTTC GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	3551	GCCGCGTCC	ACGCCATCC	C GCATCTGAC	C ACCAGCGAA	A TGGATTTTTG	
GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	0001	CGGCGCAGG	TGCGGTAGG	G CGTAGACTG	G TGGTCGCTT	r acctaaaaac	
GTAGCTCGAC CCATTATTCG CAACCGTTAA ATTGGCGGTC AGTCCGAAAG 3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG							
3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTT TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	3601	CATCGAGCT	G GGTAATAAG	C GTTGGCAAT	T TAACCGCCA	G TCAGGCTTTC	
3651 TTTCACAGAT GTGGATTGGC GATAAAAAAC AACTGCTGAC GCCGCTGCGC AAAGTGTCTA CACCTAACCG CTATTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG		GTAGCTCGA	C CCATTATTC	G CAACCGTTA	A ATTGGCGGT	C AGTCCGAAAG	
AAAGTGTCTA CACCTAACCG CTATTTTTG TTGACGACTG CGGCGACGCG 3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG							
3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	3651	TTTCACAGA'	T GTGGATTGG	C GATAAAAAA	C AACTGCTGA	C GCCGCTGCGC	
3701 GATCAGTTCA CCCGTGCACC GCTGGATAAC GACATTGGCG TAAGTGAAGC CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG		AAAGTGTCT	A CACCTAACC	G CTATTTTT	G IIGACGACT	G CGGCGACGCG	
CTAGTCAAGT GGGCACGTGG CGACCTATTG CTGTAACCGC ATTCACTTCG 3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG							
3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG	3701	GATCAGTTC	A CCCGTGCAC	C CCACCTATT	G CTGTAACCG	C ATTCACTTCG	
3751 GACCCGCATT GACCCTAACG CCTGGGTCGA ACGCTGGAAG GCGGCGGGCC CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG			- GGGCACGIG				
CTGGGCGTAA CTGGGATTGC GGACCCAGCT TGCGACCTTC CGCCGCCCGG			т сассставо	CCTGGGTCC	A ACGCTGGAA	G GCGGCGGCC	
	3/31	CTGGGCGTA	A CTGGGATTO	C GGACCCAG	T TGCGACCTI	C CGCCGCCCGG	

3801	ATTACCAGGC (CGAAGCAGCG GCTTCGTCGC	TTGTTGCAGT AACAACGTCA	GCACGGCAGA CGTGCCGTCT	TACACTTGCT ATGTGAACGA	
3851	GATGCGGTGC CTACGCCACG	TGATTACGAC ACTAATGCTG	CGCTCACGCG GCGAGTGCGC	TGGCAGCATC ACCGTCGTAG	AGGGGAAAAC TCCCCTTTTG	
3901	CTTATTTATC .	AGCCGGAAAA TCGGCCTTTT	CCTACCGGAT GGATGGCCTA	TGATGGTAGT ACTACCATCA	CCAGTTTACC	
3951	CGATTACCGT GCTAATGGCA	TGATGTTGAA ACTACAACTT	GTGGCGAGCG CACCGCTCGC	ATACACCGCA TATGTGGCGT	TCCGGCGCGG AGGCCGCGCC	
4001	ATTGGCCTGA TAACCGGACT	ACTGCCAGCT TGACGGTCGA	GGCGCAGGTA CCGCGTCCAT	GCAGAGCGGG CGTCTCGCCC	TAAACTGGCT ATTTGACCGA	
	CGGATTAGGG GCCTAATCCC	CCGCAAGAAA GGCGTTCTTT	ACTATCCCGA TGATAGGGCT	CCGCCTTACT GGCGGAATGA	GCCGCCTGTT CGGCGGACAA	
4101 1101	TTGACCGCTG AACTGGCGAC	GGATCTGCCA CCTAGACGGT	TTGTCAGACA AACAGTCTGT	TGTATACCCC ACATATGGGG	GTACGTCTTC CATGCAGAAG	
4151		TGCCAGACGC	CTGCGGGACG GACGCCCTGC	GCGCTTAACT	ATTATGGCCC TAATACCGGG	
4201	ACACCAGTGG TGTGGTCACC	CGCGGCGACT	TCCAGTTCAA AGGTCAAGTT	CATCAGCCGC	TACAGTCAAC ATGTCAGTTG	
42 51	AGCAACTGAT	GGAAACCAGC	CATCGCCATC	TGCTGCACGC ACGACGTGCC	GGAAGAAGGC CCTTCTTCCG	
4301		TATAGCTGCC	TTTCCATATO	CCCTAACCA	C CGCTGCTGAG	
	CTGGAGCCCG GACCTCGGGC	TCAGTATCGG AGTCATAGCG	GCGGAATTCCA GCCTTAAGGT	GCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	C GGTCGCTACC G CCAGCGATGG	
4401	ATTACCAGTT TAATGGTCAA	GGTCTGGTGT CCAGACCACA	CAAAAAAGA	CTGGAGGTG A GACCTCCAC	G TGGCAGCAGG C ACCGTCGTCC	
4451	GGAACCGCGC	GGCCTAGGA	T AATTAACAA' A TTAATTGTT.	A ACTGGCCAT	A TAATAGGTAG T ATTATCCATC	
	TATTCACTGA	CTAATCTAC	G TAACTAGGG	A GCTGGTTAA	C CGGTTATTT G GCCAATAAAA	
	GGTGGTATAZ	CGGCAGAAA	A CCGTTACAC	T CCCGGGCCI	A ACCTGGCCCT TGGACCGGGA	
4601	CAGAAGAAC'	r GCTCGTAAG	C TAGGGGTCT G ATCCCCAGA	A AGGGGAGAC	CG CCAAAGGAAT GC GGTTTCCTTA	
	CGTTCCAGA	C AACTTACAG	C ACTTCCTTC	G TCAAGGAGA		
4701	CTTCTGTTT	G TTGCAGACA	T CGCTGGGAZ	AA CGTCCGTC	CG GAACCCCCA GC CTTGGGGGGT	

4751	CCTGGCGACA	GGTGCCTCTG	CGGCCAAAAG	CCACGTGTAT	AAGATACACC	
		CCACGGAGAC				
						
4901	TGCAAAGGCG	CCACAACCCC	ΔΩΨΩΟΟΔΟΩΨ	TGTGAGTTGG	ATAGTTGTGG	
4001		CGTGTTGGGG				
	ACGITICCGC					
4051		T T C C C T C T C C C	man nacamam	mca a ca a cc	<u>ርር</u> ሞር እ አርር እጥ	
4851	AAAGAGTCAA	TACCGAGAGG	TCAAGCGTAT	A CHITCHITCCC	CCACTTCCTA	
	TTTCTCAGTT	TACCGAGAGG	AGIICGCAIA	AGIIGIICCC	CGACTICCIA	
4901	GCCCAGAAGG	TACCCCATTG	TATGGGATCT	GATCTGGGGC	CTCGGTGCAC	
	CGGGTCTTCC	ATGGGGTAAC	ATACCCTAGA	CTAGACCCCG	GAGCCACGTG	
4951	ATGCTTTACA	TGTGTTTAGT	CGAGGTTAAA	AAACGTCTAG	GCCCCCGAA	
	TACGAAATGT	ACACAAATCA	GCTCCAATTT	TTTGCAGATC	CGGGGGGCTT	
5001	CCACGGGGAC	GTGGTTTTCC	ТТТСАААААС	ACGATGATAA	TACCATGATT	
3001		CACCAAAAGG				
	0010000010					
E051	GAACAAGATG	CAMMCCA CCC	A CC中型C型CCC	CCCCTTCCC	тесьсьсест	
2021	GAACAAGATG	GATTGCACGC	AGGTTCTCCG	CCCCCAACCC	A CCTCTCCCA	
:::::::::::::::::::::::::::::::::::::::	CTTGTTCTAC	CTAACGTGCG	TCCAAGAGGC	CGGCGAACCC	ACCICICCGA	
	ATTCGGCTAT					
111	TAAGCCGATA	CTGACCCGTG	TTGTCTGTTA	GCCGACGAGA	CTACGGCGGC	
:= 			- 			
5T51	TGTTCCGGCT	GTCAGCGCAG	GGGCGCCCGG	TTCTTTTTGT	CAAGACCGAC	
	ACAAGGCCGA	CAGTCGCGTC	CCCGCGGGCC	AAGAAAAACA	GTTCTGGCTG	
			- 			
5201	CTCTCCCTC	CCCTCAATCA	አርጥርር አርርአር	GAGGCAGCGC	GGCTATCGTG	
5201	CACACCCCAC	CCCIGAAIGA	TGACGTCCTC	CTCCGTCGCG	CCGATAGCAC	
C.	GACAGGCOAC	000110111101	10110010010			
10		* GGGGGGG			CTTCTCTCTCTCTC	
52 51	GCTGGCCACG	ACGGGCGTTC	CTTGCGCAGC	NCACCACCTC	GTTGTCACTG CAACAGTGAC	
1.1	CGACCGGTGC	TGCCCGCAAG	GAACGCGICG	ACACGAGCIG	CARCAGIONC	
#4 #4						
	AAGCGGGAAG	GGACTGGCTG	CTATTGGGCG	AAGTGCCGGG	GCAGGATCTC	
11	TTCGCCCTTC	CCTGACCGAC	GATAACCCGC	TTCACGGCCC	CGTCCTAGAG	
5351	CTGTCATCTC	: ACCTTGCTCC	TGCCGAGAA	GTATCCATCA	TGGCTGATGC	•
	GACAGTAGAG	TGGAACGAGG	ACGGCTCTT	CATAGGTAGT	ACCGACTACG	
5401	AATGCGGCGG	CTGCATACGO	C TTGATCCGG	C TACCTGCCC#	TTCGACCACC	
	TTACGCCGCC	GACGTATGC	AACTAGGCC	G ATGGACGGGT	T AAGCTGGTGG	
5/51	AACCCAAACI	TCGCATCGA	GGAGCACGT)	A CTCGGATGG	A AGCCGGTCTT	
2421	TTCGCTTTG	r AGCGTAGCT	C GCTCGTGCA	r GAGCCTACC	TCGGCCAGAA	
					G CGCCAGCCGA	
5501	GTCGATCAG	ATGATCTGG	T CCTTCTCT	A CTCCCCCAC	C GCGGTCGGCT	
5551	ACTGTTCGC	C AGGCTCAAG	G CGCGCATGC	C CGACGGCGA	G GATCTCGTCG	
	TGACAAGCG	G TCCGAGTTC	C GCGCGTACG	G GCTGCCGCT(C CTAGAGCAGC	
5601	TGACCCATG	G CGATGCCTG	C TTGCCGAAT	A TCATGGTGG	A AAATGGCCGC	
	ACTGGGTAC	C GCTACGGAC	G AACGGCTTA		T TTTACCGGCG	
5651	TTTTCTGGA	T TCATCGACT	G TGGCCGGCT	G GGTGTGGCG	G ACCGCTATCA	
	AAAAGACCT	A AGTAGCTGA	C ACCGGCCGA	C CCACACCGC	C TGGCGATAGT	

5701	CCTGTATCGC	TTGGCTACCC AACCGATGGG	CACTATAACG	TGAAGAGCTT ACTTCTCGAA	GGCGGCGAAT CCGCCGCTTA	
5751	GGGCTGACCG	CTTCCTCGTG	CTTTACGGTA	TCGCCGCTCC AGCGGCGAGG	CGATTCGCAG GCTAAGCGTC	
5801	CGCATCGCCT GCGTAGCGGA	TCTATCGCCT AGATAGCGGA	TCTTGACGAG AGAACTGCTC	TTCTTCTGAG AAGAAGACTC	CGGGACTCTG GCCCTGAGAC	
5851	GGGTTCGCAT CCCAAGCGTA	CGATAAAATA GCTATTTTAT	AAAGATTTTA TTTCTAAAAT	TTTAGTCTCC AAATCAGAGG	TCTTTTTCCC	
5901	GGGAATGAAA CCCTTACTTT	GACCCCACCT CTGGGGTGGA	GTAGGTTTGG CATCCAAACC	CAAGCTAGCT GTTCGATCGA	TAAGTAACGC	
5951	CATTTTGCAA GTAAAACGTT	GGCATGGAAA CCGTACCTTT	AATACATAAC TTATGTATTG	TGAGAATAGA ACTCTTATCT	GAAGTTCAGA CTTCAAGTCT	
					-	
6001 \] }	TCAAGGTCAG AGTTCCAGTC	GAACAGATGG CTTGTCTACC	AACAGCTGAA TTGTCGACTT	TATGGGCCAA ATACCCGGTT	ACAGGATATC TGTCCTATAG	
6051	TGTGGTAAGC ACACCATTCG	AGTTCCTGCC TCAAGGACGG	CCGGCTCAGG GGCCGAGTCC	GCCAAGAACA CGGTTCTTGT	CTACCTTGTC	
	CTGAATATGG GACTTATACC	GCCAAACAGG CGGTTTGTCC	ATATCTGTGG TATAGACACC	TAAGCAGTTC ATTCGTCAAG	CTGCCCCGGC GACGGGGCCG	
6151	TCAGGGCCAA AGTCCCGGTT	GAACAGATGG CTTGTCTACC	TCCCCAGATG	CGGTCCAGCC	CTCAGCAGTT GAGTCGTCAA	
6201	AGATCTCTTG	CATCAGATGT GTAGTCTACA	TTCCAGGGTG	CCCCAAGGAC GGGGTTCCTG	GACTTTACTG	
62 51	CCTGTGCCTT	ATTTGAACTA TAAACTTGAT	A ACCAATCAGT TGGTTAGTCA	TCGCTTCTCG AGCGAAGAGC	CTTCTGTTCG GAAGACAAGC	
6301	CGCGCTTCTG GCGCGAAGAC	CTCCCGAGG	C TCAATAAAAC G AGTTATTTTC	AGCCCACAAC TCGGGTGTTG	CCCTCACTCG GGGAGTGAGC	
6351	GGGCGCCAGT CCCGCGGTCA	CCTCCGATTO GGAGGCTAA	G ACTGAGTCGC	CCGGGTACCC GGCCCATGGG	GTGTATCCAA CACATAGGTT	
6401	TAAACCCTCT ATTTGGGAGA	TGCAGTTGCA ACGTCAACG	A TCCGACTTG T AGGCTGAAC	T GGTCTCGCTG A CCAGAGCGAC	TTCCTTGGGA	
	CCCAGAGGA	ACTCACTAA	C TGATGGGCA	G TCGCCCCCA	TTTCATTCAT AAAGTAAGTA	
6501	GCAGCATGTA CGTCGTACA	A TCAAAATTA I AGTTTTAAT	A TTTGGTTTT T AAACCAAAA	T TTTCTTAAG A AAAGAATTC	TAAATGTAAT	
6551	AATGGCCAT: TTACCGGTA	A GTTGCATTA I CAACGTAAT	A TGAATCGGC	C AACGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	G GAGAGGCGGT C CTCTCCGCCA	
6601	TTGCGTATT AACGCATAA	G GCGCTCTTC	C GCTTCCTC	C TCACTGACT	C GCTGCGCTCG G CGACGCGAGC	

6651	GTCGTTCGGC	TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG	CGGTAATACG	
	CAGCAAGCCG	ACGCCGCTCG	CCATAGTCGA	GTGAGTTTCC	GCCATTATGC	
6701	GTTATCCACA	GAATCAGGGG	ATAACGCAGG	AAAGAACATG	TGAGCAAAAG	
	CAATAGGTGT	CTTAGTCCCC	TATTGCGTCC	TTTCTTGTAC	ACTCGTTTTC	
6751	GCCAGCAAAA					
	CGGTCGTTTT	CCGGTCCTTG	GCATTTTTCC	GGCGCAACGA	CCGCAAAAAG	
6801	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC	AAAAATCGAC	GCTCAAGTCA	
	GTATCCGAGG	CGGGGGGACT	GCTCGTAGTG	TTTTTAGCTG	CGAGTTCAGT	
6851	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG	TTTCCCCCTG	
	CTCCACCGCT	TTGGGCTGTC	CTGATATTTC	TATGGTCCGC	AAAGGGGGAC	
6901	GAAGCTCCCT					
	CTTCGAGGGA	GCACGCGAGA	GGACAAGGCT	GGGACGGCGA	ATGGCCTATG	
	CTGTCCGCCT	TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC	ATAGCTCACG	
2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	GACAGGCGGA	AAGAGGGAAG	CCCTTCGCAC	CGCGAAAGAG	TATCGAGTGC	
į∰						
7001					CTGGGCTGTG	
	GACATCCATA	GAGTCAAGCC	ACATCCAGCA	AGCGAGGTTC	GACCCGACAC	
544						
7051	TGCACGAACC	CCCCGTTCAG	CCCGACCGCT	GCGCCTTATC	CGGTAACTAT	
. 776	ACGTGCTTGG	GGGGCAAGTC	GGGCTGGCGA	CGCGGAATAG	GCCATTGATA	
7101	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	TTATCGCCAC	TGGCAGCAGC	
Ħ	GCAGAACTCA	GGTTGGGCCA	TTCTGTGCTG	AATAGCGGTG	ACCGTCGTCG	
7151	CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT	GCTACAGAGT	
5 E	GTGACCATTG	TCCTAATCGT	CTCGCTCCAT	ACATCCGCCA	CGATGTCTCA	
 − −						
7201	TCTTGAAGTG	GTGGCCTAAC	TACGGCTACA	CTAGAAGAAC	AGTATTTGGT	
£13	AGAACTTCAC	CACCGGATTO	ATGCCGATG	GATCTTCTTC	TCATAAACCA	
7251	ATCTGCGCTC	TGCTGAAGCC	AGTTACCTT	C GGAAAAAGAC	TTGGTAGCTC	
	TAGACGCGAG	ACGACTTCG	TCAATGGAA	G CCTTTTTCT(C AACCATCGAG	
7301	TTGATCCGGC	AAACAAACC	A CCGCTGGTA	G CGGTGGTTT	TTTGTTTGCA	
	AACTAGGCCG	TTTGTTTGG	r GGCGACCAT	C GCCACCAAA!	A AAACAAACGT	
7351	AGCAGCAGAT	TACGCGCAG	A AAAAAAGGA	r ctcaagaag	A TCCTTTGATC	
	TCGTCGTCTA	ATGCGCGTC	r tittitcci.	A GAGTTCTTC	r aggaaactag	
7401	TTTTCTACGO	G GGTCTGACG	C TCAGTGGAA	C GAAAACTCA	C GTTAAGGGAT	
	AAAAGATGC	CCAGACTGC	G AGTCACCTT	G CTTTTGAGT	G CAATTCCCTA	•
7451	TTTGGTCAT	G AGATTATCA	A AAAGGATCT	T CACCTAGAT	C CTTTTGCGGC	
					G GAAAACGCCG	
7501	CGCAAATCA	A TCTAAAGTA	T ATATGAGTA	A ACTTGGTCT	G ACAGTTACCA	
				T TGAACCAGA	C TGTCAATGGT	
7551	ATGCTTAAT	C AGTGAGGCA	C CTATCTCAC	C GATCTGTCT	A TTTCGTTCAT	
	TACGAATTA	G TCACTCCGT	G GATAGAGTO	G CTAGACAGA	T AAAGCAAGTA	

7601	CCATAGTTGC C					
7651	TTACCATCTG C					
7701	GGCTCCAGAT T					
7751	GAAGTGGTCC T					
7801	CGGGAAGCTA (GCCCTTCGAT (
7851 	TGCCATTGCT A ACGGTAACGA					
79 01	CATTCAGCTC (GTAAGTCGAG				TAGGGGGTAC	
7951	TTGTGCAAAA A				AACAGTCTTC	
8001 \} \}	TAAGTTGGCC ATTCAACCGG	CGTCACAATA	CACTCATGGT GTGAGTACCA	ATACCGTCGT	CTGCATAATT GACGTATTAA	
2051 11	CTCTTACTGT GAGAATGACA	CATGCCATCC GTACGGTAGG	CATTCTACGA	AAAGACACTG	ACCACTCATG	
8101	TCAACCAAGT AGTTGGTTCA	GTAAGACTCT	ATAGTGTATG TATCACATAC	GCCGCTGGCT	CAACGAGAAC	
8151 	CCCGGCGTCA GGGCCGCAGT	TATGCCCTAT	TATGGCGCGG	TGTATCGTCT	TGAAATTTTC	
8201	ACGAGTAGTA	ACCTTTTGCA	AGAAGCCCCG	CTTTTGAGAG	AAGGATCTTA TTCCTAGAAT	
		CTAGGTCAAG	CTACATTGGG	TGAGCACGTG	GGTTGACTAG	
		AAATGAAAGT	GGTCGCAAAC	ACCCACTCGT	TTTTGTCCTT	
		GCGTTTTTTC	CCTTATTCC	C GCTGTGCCTT	TACAACTTAT	
		rəaaaaaəa 	TATAATAAC	TCGTAAATA(TCCCAATAAC	-
		CCTATGTATA	A AACTTACAT	A AATCTTTTT	A TTTGTTTATC	
8501	GGGTTCCGCG CCCAAGGCGC					

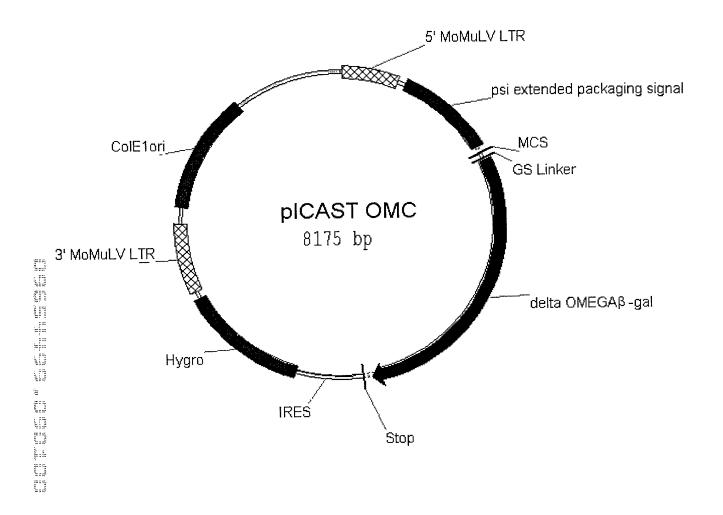


Figure 12A

1	CTGCAGCCTG	AATATGGGCC	AAACAGGATA	TCTGTGGTAA	GCAGTTCCTG	
	GACGTCGGAC	TTATACCCGG	TTTGTCCTAT	AGACACCATT	CGTCAAGGAC	
51	CCCCGGCTCA	GGGCCAAGAA	CAGATGGAAC	AGCTGAATAT	GGGCCAAACA	
	GGGGCCGAGT	CCCGGTTCTT	GTCTACCTTG	TCGACTTATA	CCCGGTTTGT	
101	GGATATCTGT	GGTAAGCAGT	TCCTGCCCCG	GCTCAGGGCC	AAGAACAGAT	
	CCTATAGACA	CCATTCGTCA	AGGACGGGGC	CGAGTCCCGG	TTCTTGTCTA	
151	GGTCCCCAGA	TGCGGTCCAG	CCCTCAGCAG	TTTCTAGAGA	ACCATCAGAT	
	CCAGGGGTCT	ACGCCAGGTC	GGGAGTCGTC	AAAGATCTCT	TGGTAGTCTA	
				-		
201	GTTTCCAGGG	TGCCCCAAGG	ACCTGAAATG	ACCCTGTGCC	TTATTTGAAC	
	CAAAGGTCCC	ACGGGGTTCC	TGGACTTTAC	TGGGACACGG	AATAAACTTG	
251	TAACCAATCA	GTTCGCTTCT	CGCTTCTGTT	CGCGCGCTTC	TGCTCCCCGA	
		CAAGCGAAGA				
301	GCTCAATAAA	AGAGCCCACA	ACCCCTCACT	CGGGGCGCCA	GTCCTCCGAT	
M.	CGAGTTATTT	TCTCGGGTGT	TGGGGAGTGA	GCCCCGCGGT	CAGGAGGCTA	
TI						
351	TGACTGAGTC	GCCCGGGTAC	CCGTGTATCC	AATAAACCCT	CTTGCAGTTG	
	ACTGACTCAG	CGGGCCCATG	GGCACATAGG	TTATTTGGGA	GAACGTCAAC	
: ==						
4 -01	CATCCGACTT	GTGGTCTCGC	TGTTCCTTGG	GAGGGTCTCC	TCTGAGTGAT	
401	GTAGGCTGAA	CACCAGAGCG	ACAAGGAACC	CTCCCAGAGG	AGACTCACTA	
.77						
<u>.</u> 451	TGACTACCCG	TCAGCGGGGG	TCTTTCATTT	GGGGGCTCGT	CCGGGATCGG	:
5.£	ACTGATGGGC	AGTCGCCCCC	AGAAAGTAAA	CCCCGAGCA	GGCCCTAGCC	
W	. 					
501	GAGACCCCTG	CCCAGGGACC	ACCGACCCAC	CACCGGGAGG	CAAGCTGGCC	
ļ.ā	CTCTGGGGAC	GGGTCCCTGG	TGGCTGGGTG	GTGGCCCTCC	GTTCGACCGG	
<u> </u>						
5 51	AGCAACTTAT	CTGTGTCTGT	CCGATTGTCT	AGTGTCTATO	ACTGATTTTA	
###	TCGTTGAATA	GACACAGACA	GGCTAACAGA	TCACAGATAC	TGACTAAAAT	
601	TGCGCCTGCG	TCGGTACTAG	TTAGCTAACT	AGCTCTGTAT	T CTGGCGGACC	
	ACGCGGACGC	AGCCATGATC	AATCGATTGA	TCGAGACATA	A GACCGCCTGG	
651	CGTGGTGGA	CTGACGAGTT	CTGAACACC	C GGCCGCAAC	C CTGGGAGACG	
	GCACCACCTT	GACTGCTCAA	GACTTGTGG	CCGGCGTTG	G GACCCTCTGC	
701	TCCCAGGGA	TTTGGGGGCC	GTTTTTGTG	G CCCGACCTG	A GGAAGGGAGT	
	AGGGTCCCTC	AAACCCCCG	CAAAAACAC	C GGGCTGGAC	r ccttccctca	
751	CGATGTGGA	A TCCGACCCC	G TCAGGATAT	G TGGTTCTGG	T AGGAGACGAG	
					A TCCTCTGCTC	
801	AACCTAAAA	C AGTTCCCGC	C TCCGTCTGA	A TTTTTGCTT	T CGGTTTGGAA	-
					A GCCAAACCTT	
851	CCGAAGCCG	C GCGTCTTGT	C TGCTGCAGC	A TCGTTCTGT	G TTGTCTCTGT	
	GGCTTCGGC	G CGCAGAACA	G ACGACGTCG	T AGCAAGACA	.C AACAGAGACA	
901	CTGACTGTG	T TTCTGTATT	T GTCTGAAAA	T TAGGGCCAG	A CTGTTACCAC	
	GACTGACAC	A AAGACATAA	A CAGACTTT	A ATCCCGGTC	T GACAATGGTG	

951	TCCCTTAAGT AGGGAATTCA	TTGACCTTAG AACTGGAATC	GTAACTGGAA CATTGACCTT	AGATGTCGAG TCTACAGCTC	CGGCTCGCTC GCCGAGCGAG	
1001	ACAACCAGTC TGTTGGTCAG	GGTAGATGTC CCATCTACAG	AAGAAGAGAC TTCTTCTCTG	GTTGGGTTAC CAACCCAATG	CTTCTGCTCT GAAGACGAGA	
1051	GCAGAATGGC CGTCTTACCG	CAACCTTTAA GTTGGAAATT	CGTCGGATGG GCAGCCTACC	CCGCGAGACG GGCGCTCTGC	GCACCTTTAA CGTGGAAATT	
1101	CCGAGACCTC GGCTCTGGAG	ATCACCCAGG TAGTGGGTCC	TTAAGATCAA AATTCTAGTT	GGTCTTTTCA CCAGAAAAGT	CCTGGCCCGC GGACCGGGCG	
1151	ATGGACACCC TACCTGTGGG	AGACCAGGTC TCTGGTCCAG	CCCTACATCG GGGATGTAGC	TGACCTGGGA ACTGGACCCT	AGCCTTGGCT TCGGAACCGA	
1201	TTTGACCCCC AAACTGGGGG	CTCCCTGGGT GAGGGACCCA	CAAGCCCTTT GTTCGGGAAA	GTACACCCTA CATGTGGGAT	AGCCTCCGCC TCGGAGGCGG	
1251	TCCTCTTCCT AGGAGAAGGA	CCATCCGCCC GGTAGGCGGG	CGTCTCTCCC GCAGAGAGGG	CCTTGAACCT GGAACTTGGA	CCTCGTTCGA GGAGCAAGCT	
1301 13	CCCCGCCTCG GGGGCGGAGC	ATCCTCCCTT TAGGAGGGAA	TATCCAGCCC ATAGGTCGGG	TCACTCCTTC AGTGAGGAAG	TCTAGGCGCC AGATCCGCGG	
1351	GGCCGCTCTA CCGGCGAGAT	GCCCATTAAT	ACGACTCACT TGCTGAGTGA	ATAGGGCGAT TATCCCGCTA	TCGAATCAGG AGCTTAGTCC	
1401 2	CCTTGGCGCG GGAACCGCGC	CCGGATCCTT GGCCTAGGAA	AATTAAGCGC TTAATTCGCG	AATTGGGAGG TTAACCCTCC	TGGCGGTAGC ACCGCCATCG	
1451 4	CTCGAGATGG GAGCTCTACC	GCGTGATTAC CGCACTAATG	GGATTCACTG	GCCGTCGTTT CGGCAGCAAA	TACAACGTCG ATGTTGCAGC	
15 01	TGACTGGGAA ACTGACCCTT	AACCCTGGCG	TTACCCAACT AATGGGTTGA	TAATCGCCTT ATTAGCGGAA	GCAGCACATC CGTCGTGTAG	
1551	CCCCTTTCG(GGGGAAAGC(C CAGCTGGCGT G GTCGACCGCF	AATAGCGAAC	AGGCCCGCAC TCCGGGCGTG	C CGATCGCCCT G GCTAGCGGGA	
1601	TCCCAACAGT AGGGTTGTCA	TACGCAGCCT A ATGCGTCGGA	GAATGGCGAM CTTACCGCT	A TGGCGCTTT(T ACCGCGAAA(C CCTGGTTTCC C GGACCAAAGG	
	CCGTGGTCT	r cgccacggc	TTTCGACCG	A CCTCACGCTA		
	GGCTATGAC	A GCAGCAGGG	G AGTTTGACC	G TCTACGTGC		
	GGGTAGATG	T GGTTGCACT	G GATAGGGTA 	A TGCCAGTTA		
	AGGGTGCCT	C TTAGGCTGC	C CAACAATGA 	G CGAGTGTAA		
1851	TTTCGACCG	A TGTCCTTCC	G GTCTGCGCT	T AATAAAAAC	A TGGCGTTAAC T ACCGCAATTG	

1901	TCGGCGTTTC AGCCGCAAAG	ATCTGTGGTG TAGACACCAC	CAACGGGCGC GTTGCCCGCG	TGGGTCGGTT ACCCAGCCAA	ACGGCCAGGA TGCCGGTCCT	
1951	CAGTCGTTTG GTCAGCAAAC	CCGTCTGAAT GGCAGACTTA	TTGACCTGAG AACTGGACTC	CGCATTTTTA GCGTAAAAAT	CGCGCCGGAG GCGCGGCCTC	
2001	AAAACCGCCT TTTTGGCGGA	CGCGGTGATG GCGCCACTAC	GTGCTGCGCT CACGACGCGA	GGAGTGACGG CCTCACTGCC	GTCAATAGAC	
2051	GAAGATCAGG CTTCTAGTCC	TATACACCGC	GATGAGCGGC CTACTCGCCG	TAAAAGGCAC	TGCAGAGCAA	
2101	GCTGCATAAA CGACGTATTT	CCGACTACAC GGCTGATGTG	AAATCAGCGA TTTAGTCGCT	TTTCCATGTT AAAGGTACAA	GCCACTCGCT CGGTGAGCGA	
2151	TTAATGATGA AATTACTACT	AAAGTCGGCG	GCTGTACTGG CGACATGACC	AGGCTGAAGT TCCGACTTCA	TCAGATGTGC AGTCTACACG	
2 20 1	GGCGAGTTGC CCGCTCAACG	GTGACTACCT CACTGATGGA	ACGGGTAACA TGCCCATTGT	GTTTCTTTAT CAAAGAAATA	CCGTCCCACT	
2 25 1		GCCAGCGGCA CGGTCGCCGT	CCGCGCCTTT GGCGCGGAAA	CGGCGGTGAA GCCGCCACTT	ATTATCGATG TAATAGCTAC	
2301	AGCGTGGTGG TCGCACCACC	TTATGCCGAT AATACGGCTA	CGCGTCACAC GCGCAGTGTG	TACGTCTGAA ATGCAGACTT	CGTCGAAAAC GCAGCTTTTG	
		CCTCGCGGCT	TTAGGGCTTA	GAGATAGCAC	GCCACCAACT	
2401 =		CGGCTGCCGT	GCGACTAACT	TCGTCTTCGG	ACGCTACAGC	
2451 	CAAAGGCGCT	CCACGCCTAA	CTTTTACCAG	ACGACGACGA	GAACGGCAAG CTTGCCGTTC	
		AAGCTCCGCA	ATTGGCAGTG	CTCGTAGTAG	GAGACGTACC	
2551	AGTCCAGTAC	CTACTCGTCT	GCTACCACG	CCTATAGGAC	G CTGATGAAGC C GACTACTTCG	
	TCTTGTTGAA	ATTGCGGCAC	GCGACAAGC	TAATAGGCT	A CCATCCGCTG GGTAGGCGAC ATGAAGCCAA	·
	ACCATGTGCC	ACACGCTGG	C GATGCCGGA	C ATACACCACO	C TACTTCGGTT	
	ATAACTTTG	GTGCCGTAC	C ACGGTTACT	T AGCAGACTG	G CTACTAGGCG	
	CGACCGATG	G CCGCTACTC	G CTTGCGCAT	T GCGCTTACC	A CGTCGCGCTA	
2001	GCATTAGTG	G GCTCACACT	A GTAGACCAG	C GACCCCTTA	C TTAGTCCGGT	

2851	CGGCGCTAAT (GCCGCGATTA (
2901	CCCGCCGGT (
2951	ATTATTTGCC TAATAAACGG					
3001	TGTGCCGAAA ACACGGCTTT					
3051	GCCCGCTGAT CGGGCGACTA	CCTTTGCGAA GGAAACGCTT	TACGCCCACG ATGCGGGTGC	CGATGGGTAA GCTACCCATT	CAGTCTTGGC GTCAGAACCG	
3101	GGTTTCGCTA CCAAAGCGAT	AATACTGGCA TTATGACCGT	GGCGTTTCGT CCGCAAAGCA	CAGTATCCCC GTCATAGGGG	GTTTACAGGG CAAATGTCCC	
3151	CGGCTTCGTC GCCGAAGCAG		TGGATCAGTC ACCTAGTCAG			
	ACGGCAACCC TGCCGTTGGG	CACCAGCCGA	TACGGCGGTG ATGCCGCCAC	TAAAACCGCT	ATGCGGCTTG	
3251		AGACATACTT	CGGTCTGGTC GCCAGACCAG	AAACGGCTGG		
3301 -			AACACCAGCA TTGTGGTCGT			
3351			GTGACCAGCG CACTGGTCGC		GGCAGTATCG	
3.401	GATAACGAGC CTATTGCTCG	TCCTGCACTG AGGACGTGAC	GATGGTGGCG CTACCACCGC	CTGGATGGTA GACCTACCAT	AGCCGCTGGC TCGGCGACCG	
3451	AAGCGGTGAA TTCGCCACTT		ATGTCGCTCC TACAGCGAGG			
		TGATGGCGTC	GGCCTCTCGC	GGCCCGTTGA		
		ACGTTGGCTT	GCGCTGGCGT	ACCAGTCTTC	GGCCCGTGTA	
3601		CAGCAGTGGC GTCGTCACCG	CAGACCGCCT	TTTGGAGTC	A CACTGCGAGG	
3651		CCACGCCATC GGTGCGGTAG	G GGCGTAGACT	CCACCAGCGA GGTGGTCGCA	A AATGGATTTT T TTACCTAAAA	
3701	ACGTAGCTCG	ACCCATTATI	CGCAACCGTI	AAATTGGCG		
3751					G ACGCCGCTGC C TGCGGCGACG	

3801	GCGATCAGTT CGCTAGTCAA					
3851	GAAGACCTAG CTTCTGGATC					
3901	TAAGTGACTG ATTCACTGAC		TTTCGACTAG AAAGCTGATC			
3951			AGAAAACCGT			
4001	GGCCCTGTCT	TCTTGACGAG				
4051	AGGAATGCAA TCCTTACGTT		ATGTCGTGAA TACAGCACTT			
4101	CTTCTTGAAG GAAGAACTTC		TCTGTAGCGA AGACATCGCT		CGTCGCCTTG	
4151 UT	CCCCACCTG GGGGGTGGAC		CCTCTGCGGC GGAGACGCCG			
4201			AACCCCAGTG TTGGGGTCAC			
4251			CTCTCCTCAA GAGAGGAGTT			
		TCTTCCATGG	CCATTGTATG GGTAACATAC	CCTAGACTAG	TGGGGCCTCG ACCCCGGAGC	
4351	GTGCACATGC CACGTGTACG	TTTACATGTG AAATGTACAC	TTTAGTCGAG AAATCAGCTC	GTTAAAAAAC CAATTTTTG	GTCTAGGCCC CAGATCCGGG	
4401	CCCGAACCAC GGGCTTGGTG	GGGGACGTGG CCCCTGCACC	TTTTCCTTTG AAAAGGAAAC	AAAAACACGA TTTTTGTGCT	TGATAATACC ACTATTATGG	
	ATGAAAAAGC TACTTTTTCG	GACTTGAGTG	GCGCTGCAGA	CAGCTCTTCA	TTCTGATCGA AAGACTAGCT	
4501		TCGCAGAGGC	TGGACTACGI	CGAGAGCCTC	CCGCTTCTTA	
	GAGCACGAAA	GTCGAAGCTA	A CATCCTCCC	CACCTATACA	CCTGCGGGTA GGACGCCCAT	
4601	TTATCGACGC	GGCTACCAA		A GCAATACAA	TATCGGCACTT A TAGCCGTGAA	
4651	ACGTAGCCGC	G CGCGAGGGC	r aaggeette	A CGAACTGTAA		
4701	CGCTCTCGG	A CTGGATAAC	G TAGAGGGCG	G CACGTGTCC	G TGTCACGTTG C ACAGTGCAAC	

4751	CAAGACCTGC GTTCTGGACG		ACTGCCCGCT TGACGGGCGA			
4801	GGCCATGGAT CCGGTACCTA		CGGCCGATCT GCCGGCTAGA			
4851	GCCCATTCGG CGGGTAAGCC		ATCGGTCAAT TAGCCAGTTA		CGCACTAAAG	
4901	ATATGCGCGA TATACGCGCT		CCATGTGTAT GGTACACATA	GTGACCGTTT	GACACTACCT	
4951	CGACACCGTC GCTGTGGCAG		TCGCGCAGGC AGCGCGTCCG			
5001	GGGCCGAGGA CCCGGCTCCT		GTCCGGCACC CAGGCCGTGG			
5051		AGGACTGCCT	CAATGGCCGC GTTACCGGCG	TATTGTCGCC	AGTAACTGAC	
5101 11	GAGCGAGGCG CTCGCTCCGC		ATTCCCAATA TAAGGGTTAT			
51 51			TGTATGGAGC ACATACCTCG		CTACTTCGAG GATGAAGCTC	
5201 					CGTATATGCT GCATATACGA	
	CCGCATTGGT GGCGTAACCA				CCGTTAAAGC	
5 301					CCGATCCGGA GGCTAGGCCT	
					G CGGCCGTCTG C GCCGGCAGAC	
5401					C CGACGCCCCA G GCTGCGGGGT	
	CGTGAGCAGG	CTCCCGTTTC	CTTATCTCAT	CTACGGCTGC		
	CTATTTTATT	TTCTAAAATA	A AATCAGAGGI	CTTTTTCCC		
	TGGGGTGGAC	ATCCAAACCC	TTCGATCGA	TTCATTGCG	C ATTTTGCAAG G TAAAACGTTC	·
	CGTACCTTT	TATGTATTG	A CTCTTATCTO	TTCAAGTCT		
5651	TTGTCTACCT	TGTCGACTT	A TACCCGGTT	r gtcctatag	T GTGGTAAGCA A CACCATTCGT	

5701	GTTCCTGCCC CAAGGACGGG	CGGCTCAGGG GCCGAGTCCC	CCAAGAACAG GGTTCTTGTC	ATGGAACAGC TACCTTGTCG	TGAATATGGG ACTTATACCC	
	CCAAACAGGA GGTTTGTCCT	ATAGACACCA	AAGCAGTTCC TTCGTCAAGG	TGCCCCGGCT ACGGGGCCGA	CAGGGCCAAG GTCCCGGTTC	
	AACAGATGGT TTGTCTACCA	CCCCAGATGC	GGTCCAGCCC CCAGGTCGGG	TCAGCAGTTT AGTCGTCAAA	GATCTCTTGG	
5851	ATCAGATGTT TAGTCTACAA	TCCAGGGTGC AGGTCCCACG	GGGTTCCTGG	ACTTTACTGG	GACACGGAAT	
5901	TTTGAACTAA AAACTTGATT	CCAATCAGTT GGTTAGTCAA	CGCTTCTCGC GCGAAGAGCG	TTCTGTTCGC AAGACAAGCG	GCGCTTCTGC CGCGAAGACG	
5951	TCCCCGAGCT AGGGGCTCGA	CAATAAAAGA GTTATTTTCT	GCCCACAACC CGGGTGTTGG	CCTCACTCGG GGAGTGAGCC	CCGCGGTCAG	
60 01	CTCCGATTGA GAGGCTAACT	CTGAGTCGCC GACTCAGCGG	CGGGTACCCG GCCCATGGGC	TGTATCCAAT ACATAGGTTA	TTTGGGAGAA	
60 51	GCAGTTGCAT CGTCAACGTA	CCGACTTGTG GGCTGAACAC	GTCTCGCTGT CAGAGCGACA	TCCTTGGGAG AGGAACCCTC	GGTCTCCTCT CCAGAGGAGA	
6101 11	GAGTGATTGA CTCACTAACT	CTACCCGTCA GATGGGCAGT	GCGGGGGTCT CGCCCCAGA	TTCATTCATG AAGTAAGTAC	CAGCATGTAT GTCGTACATA	
6151	GTTTTAATTA	AACCAAAAAA	AAGAATTCAT 	AAATGTAATT		
6201	AACGTAATTA	CTTAGCCGGT	TGCGCGCCCC	TCTCCGCCAA		
<u>6</u> 251		GAAGGAGCGA	GTGACTGAGC	GACGCGAGCC	AGCAAGCCGA	
		CATAGTCGAG	TGAGTTTCC	CCATTATGCC	AATAGGTGTC	
		ATTGCGTCCI	TTCTTGTAC	A CTCGTTTTCC	GGTCGTTTTC	
		CATTTTTCC	GCGCAACGA	C CGCAAAAAG(TATCCGAGGC	
	GGGGGGACTG	CTCGTAGTG	TTTTAGCTG	C GAGTTCAGT		
	TGGGCTGTCC	TGATATTTC	r atggtccgc.	A AAGGGGGAC		
	CACGCGAGAC	G GACAAGGCT	G GGACGGCGA	A TGGCCTATG		
6601	AGAGGGAAG	C CCTTCGCAC	C GCGAAAGAG	T ATCGAGTGC	C TGTAGGTATC G ACATCCATAG	

6651		GTAGGTCGTT CATCCAGCAA				
6701	CCCGTTCAGC GGGCAAGTCG	CCGACCGCTG GGCTGGCGAC				
6751	CAACCCGGTA GTTGGGCCAT	AGACACGACT TCTGTGCTGA			TGACCATTGT	·
6801	GGATTAGCAG CCTAATCGTC	AGCGAGGTAT TCGCTCCATA				
6851		ACGGCTACAC TGCCGATGTG				
6901	GCTGAAGCCA CGACTTCGGT	GTTACCTTCG CAATGGAAGC				
6951	AACAAACCAC TTGTTTGGTG	GCGACCATCG		AACAAACGTT	CGTCGTCTAA	
7001	ACGCGCAGAA TGCGCGTCTT		TCAAGAAGAT	CCTTTGATCT	TTTCTACGGG	
7051	CAGACTGCGA	CAGTGGAACG GTCACCTTGC	TTTTGAGTGC		TTGGTCATGA	
7101		TTCCTAGAAG	TGGATCTAGG	AAAATTTAAT	TTTTACTTCA	
7151	TTGCGGCCGC AACGCCGGCG	AAATCAATCT TTTAGTTAGA	AAAGTATATA	TGAGTAAACT	TGGTCTGACA	
7 201	GTTACCAATG	CTTAATCAGT GAATTAGTCA				
7251	CGTTCATCCA GCAAGTAGGT	TAGTTGCCTG ATCAACGGAC				
7301	GGAGGGCTTA CCTCCCGAAT	CCATCTGGCC GGTAGACCGG				
		AGGTCTAAAT	AGTCGTTATT	TGGTCGGTCG	GCCTTCCCGG	
7401		CACCAGGACG	TTGAAATAGG	CGGAGGTAGG	TCAGATAATT	
7451	AACAACGGCC	CTTCGATCTC	ATTCATCAAG	CGGTCAATTA		
7501	ACGTTGTTGC TGCAACAACG	CATTGCTACA GTAACGATGT	GGCATCGTGG CCGTAGCACC	TGTCACGCTC ACAGTGCGAG	GTCGTTTGGT CAGCAAACCA	
	TACCGAAGTA	AGTCGAGGCC	AAGGGTTGCT	AGTTCCGCTC	G TTACATGATC C AATGTACTAG	

7601	CCCCATGTTG GGGGTACAAC	TGCAAAAAAG ACGTTTTTTC	CGGTTAGCTC GCCAATCGAG	CTTCGGTCCT GAAGCCAGGA	CCGATCGTTG GGCTAGCAAC	
7651	TCAGAAGTAA AGTCTTCATT	GTTGGCCGCA CAACCGGCGT	GTGTTATCAC CACAATAGTG	TCATGGTTAT AGTACCAATA	GGCAGCACTG CCGTCGTGAC	
7701	CATAATTCTC GTATTAAGAG	TTACTGTCAT AATGACAGTA	GCCATCCGTA CGGTAGGCAT	AGATGCTTTT TCTACGAAAA	CTGTGACTGG GACACTGACC	
7751	TGAGTACTCA ACTCATGAGT	ACCAAGTCAT TGGTTCAGTA	TCTGAGAATA AGACTCTTAT	GTGTATGCGG CACATACGCC	CGACCGAGTT GCTGGCTCAA	
7801	GCTCTTGCCC CGAGAACGGG	GGCGTCAATA CCGCAGTTAT	CGGGATAATA GCCCTATTAT	CCGCGCCACA GGCGCGGTGT	TAGCAGAACT ATCGTCTTGA	
7851	TTAAAAGTGC AATTTTCACG	TCATCATTGG AGTAGTAACC	AAAACGTTCT TTTTGCAAGA	TCGGGGCGAA AGCCCCGCTT	AACTCTCAAG TTGAGAGTTC	
7901 45	GATCTTACCG CTAGAATGGC	CTGTTGAGAT GACAACTCTA	CCAGTTCGAT GGTCAAGCTA	GTAACCCACT CATTGGGTGA	CGTGCACCCA GCACGTGGGT	
7951	ACTGATCTTC TGACTAGAAG	AGCATCTTTT TCGTAGAAAA	ACTTTCACCA TGAAAGTGGT	GCGTTTCTGG CGCAAAGACC	GTGAGCAAAA CACTCGTTTT	
8001	ACAGGAAGGC TGTCCTTCCG	AAAATGCCGC TTTTACGGCG	AAAAAAGGGA TTTTTTCCCT	ATAAGGGCGA TATTCCCGCT	CACGGAAATG GTGCCTTTAC	
8051	TTGAATACTC AACTTATGAG	ATACTCTTCC TATGAGAAGG	TTTTTCAATA AAAAAGTTAT	TTATTGAAGC AATAACTTCG	ATTTATCAGG TAAATAGTCC	
		CATGAGCGGA GTACTCGCCT	TACATATTTC ATGTATAAAC	TTACATAAAT	CTTTTTATTT	
8151	CAAATAGGGG GTTTATCCCC	TTCCGCGCAC	TAAAG			

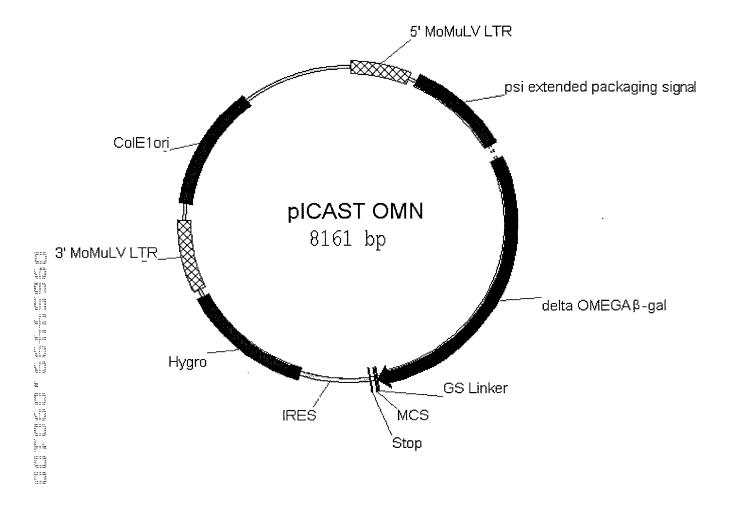


Figure 13A

1	CTGCAGCCTG GACGTCGGAC		AAACAGGATA TTTGTCCTAT			
51	CCCCGGCTCA GGGGCCGAGT		CAGATGGAAC GTCTACCTTG			
101	GGATATCTGT CCTATAGACA				AAGAACAGAT TTCTTGTCTA	
151	GGTCCCCAGA CCAGGGGTCT		CCCTCAGCAG GGGAGTCGTC			
201	GTTTCCAGGG CAAAGGTCCC				TTATTTGAAC AATAAACTTG	
251	TAACCAATCA ATTGGTTAGT				TGCTCCCCGA ACGAGGGGCT	
301	GCTCAATAAA CGAGTTATTT		TGGGGAGTGA		CAGGAGGCTA	
351 351					CTTGCAGTTG GAACGTCAAC	
401					TCTGAGTGAT AGACTCACTA	
1451 151					CCGGGATCGG GGCCCTAGCC	
## 5 01	GAGACCCCTG CTCTGGGGAC				CAAGCTGGCC	
551					ACTGATTTTA TGACTAAAAT	,
601					CTGGCGGACC GACCGCCTGG	
651					CTGGGAGACG GACCCTCTGC	
701	AGGGTCCCTG	AAACCCCCGG	CAAAAACACC	GGGCTGGACT	GGAAGGGAGT CCTTCCCTCA	
		AGGCTGGGGC	AGTCCTATAC	ACCAAGACCA	AGGAGACGAG ATCCTCTGCTC	
801	TTGGATTTTG	TCAAGGGCGG	AGGCAGACTI	AAAAACGAAA	CGGTTTGGAA GCCAAACCTT	
851	GGCTTCGGCG	CGCAGAACAG	ACGACGTCG1	AGCAAGACA	TTGTCTCTGT AACAGAGACA	
	GACTGACACA	AAGACATAAA	A CAGACTTTT	A ATCCCGGTC	A CTGTTACCAC F GACAATGGTG	

951		TTGACCTTAG AACTGGAATC				
1001		GGTAGATGTC CCATCTACAG				
1051		CAACCTTTAA GTTGGAAATT				
1101		ATCACCCAGG TAGTGGGTCC				
1151		AGACCAGGTC TCTGGTCCAG				
1201		CTCCCTGGGT GAGGGACCCA				
		CCATCCGCCC GGTAGGCGGG			GGAGCAAGCT	
1301 		ATCCTCCCTT TAGGAGGGAA				
#351 #3	CCGGCGAGAT	GCCCATTAAT CGGGTAATTA				
13 01	TGCACCATCA	TCATCATCAC AGTAGTAGTG				
1451	CTGGATGAGC	AGATGGGCGT TCTACCCGCA	CTAATGCCTA	AGTGACCGGC	AGCAAAATGT	
1501	TGCAGCACTG	TGGGAAAACC ACCCTTTTGG				
	GTGTAGGGG		ACCGCATTAT	CGCTTCTCCG	GGCGTGGCTA	
	GCGGGAAGGG	TTGTCAATGC	GTCGGACTTA	CCGCTTACCG		
	CAAAGGCCGT	GGTCTTCGCC	ACGGCCTTTC	GACCGACCTC		
	GACTCCGGCT	ATGACAGCAG	CAGGGGAGTT	TGACCGTCTA		·
	CTACGCGGGT	AGATGTGGTI	GCACTGGATA	GGGTAATGCC	G TCAATCCGCC C AGTTAGGCGG C ACATTTAATG	·
	CAAACAAGGG	G TGCCTCTTAG	G GCTGCCCAAC	C AATGAGCGAC	G TGTAAATTAC	
1821	AACTACTTT	C GACCGATGTO	CTTCCGGTCT	GCGCTTAATA	r tittgatggc A aaaactaccg	

1901	GTTAACTCGG CAATTGAGCC		GTGGTGCAAC CACCACGTTG			
1951			CTGAATTTGA GACTTAAACT			
2001	CCGGAGAAAA GGCCTCTTTT		GTGATGGTGC CACTACCACG			
2051	TATCTGGAAG ATAGACCTTC		GTGGCGGATG CACCGCCTAC		AGGCACTGCA	
2101	CTCGTTGCTG GAGCAACGAC		CTACACAAAT GATGTGTTTA			
2151	CTCGCTTTAA GAGCGAAATT		AGCCGCGCTG TCGGCGCGAC			
# F			GATGGATGCC		GAAATACCGT	
2251			GCGGCACCGC CGCCGTGGCG			
2301			GCCGATCGCG CGGCTAGCGC			
2351	CTTTTGGGCT	TTGACACCTC	CGCCGAAATC GCGGCTTTAG	GGCTTAGAGA	TAGCACGCCA	
		GTGTGGCGGC	TGCCGTGCGA	CTAACTTCGT	CTTCGGACGC	
2451	ATGTCGGTTT TACAGCCAAA		CGGATTGAAA GCCTAACTTT			
2501	CCGTTCGGCA	ACGACTAAGC	AGGCGTTAAC TCCGCAATTG	GCAGTGCTCG	TAGTAGGAGA	
		CAGTACCTAC	TCGTCTGCTA	CCACGTCCTA	TAGGACGACT	
		GTTGAAATTG	CGGCACGCGA	CAAGCGTAAT	AGGCTTGGTA	
	GGCGACACCA	TGTGCGACAC	GCTGGCGATG	CCGGACATAC		
	TCGGTTATAA	CTTTGGGTGC	CGTACCACGG	TTACTTAGCA	CTGACCGATG GACTGGCTAC	
	TAGGCGCGAC	CGATGGCCGC	TACTCGCTTG	CGCATTGCGC		
∠801 	GCGCTAGCAT	TAGTGGGCTC	ACACTAGTAG	ACCAGCGACC	GGAATGAATC CCTTACTTAG	

2851	AGGCCACGGC TCCGGTGCCG	GCTAATCACG CGATTAGTGC				
2901	ATCCTTCCCG TAGGAAGGGC	CCCGGTGCAG GGGCCACGTC				
2951	ACCGATATTA TGGCTATAAT	TTTGCCCGAT AAACGGGCTA				
3001	CCCGGCTGTG GGGCCGACAC	CCGAAATGGT GGCTTTACCA				
3051	AGACGCGCCC TCTGCGCGGG	GCTGATCCTT CGACTAGGAA				
3101	CTTGGCGGTT GAACCGCCAA	TCGCTAAATA AGCGATTTAT				
 3151	ACAGGGCGGC	 ттсстстссс		TCAGTCGCTG	ATTA A ATATC	
111		AAGCAGACCC			TAATTTATAC	
3201 []]	ATGAAAACGG TACTTTTGCC	CAACCCGTGG GTTGGGCACC			TGGCGATACG	
₫₩ 32 51	CCGAACGATC	GCCAGTTCTG	TATGAACGGT	СТССТСТТТС	CCGACCGCAC	
		CGGTCAAGAC				
3301		GCGCTGACGG CGCGACTGCC				
3351		CGGGCAAACC GCCCGTTTGG				
3401		ACGAGCTCCT TGCTCGAGGA				
3451	GCTGGCAAGC CGACCGTTCG	GGTGAAGTGC CCACTTCACG				
3501		GCCTGAACTA CGGACTTGAT				
3551					CAGAAGCCGG	-
3601	CGTGTAGTCG	CGGACCGTCG	TCACCGCAGA	CCGCCTTTTG		
3651	CGCTCCCCGC GCGAGGGGCG	GCGCAGGGTG	GCCATCCCGC CGGTAGGGCG	ATCTGACCAC TAGACTGGTG	CAGCGAAATG GTCGCTTTAC	
3701	GATTTTTGCA	TCGAGCTGGG	TAATAAGCGT	TGGCAATTTA	ACCGCCAGTC TGGCGGTCAG	
	AGGCTTTCTT	TCACAGATGT	GGATTGGCGA	TAAAAAACAA	A CTGCTGACGC T GACGACTGCG	

3801			CGTGTCGATA GCACAGCTAT			
3851			CTTAATTAAC GAATTAATTG			
3901			TGCATTTCGA ACGTAAAGCT			
3951	GCCAATAAAA	GGTGGTATAA	GCCGTCTTTT CGGCAGAAAA	CCGTTACACT		
4001			CGAGCATTCC GCTCGTAAGG			
4051			TTGAATGTCG AACTTACAGC			
4101			AACGTCTGTA TTGCAGACAT			
			GGTGCCTCTG CCACGGAGAC			
4 201			GCACAACCCC CGTGTTGGGG		TGTGAGTTGG ACACTCAACC	
4251					TCAACAAGGG AGTTGTTCCC	
43 01	CGACTTCCTA	CGGGTCTTCC		ATACCCTAGA	GATCTGGGGC CTAGACCCCG	
4351					AAACGTCTAG TTTGCAGATC	
4401					ACGATGATAA TGCTACTATT	
4451					AAGTTTCTGA TTCAAAGACT	·
4501	AGCTTTTCAA	GCTGTCGCAG	AGGCTGGACT	ACGTCGAGAG	GGAGGGCGAA CCTCCCGCTT	
4551	CTTAGAGCAC	GAAAGTCGAA		CCCGCACCTA	ATGTCCTGCG TACAGGACGC	
4601					GTTTATCGGC CAAATAGCCG	
4651	TGAAACGTAG	CCGGCGCGAG	GGCTAAGGCC	TTCACGAACT	CATTGGGGAA GTAACCCCTT	
4701	AAATCGCTCT	CGGACTGGAT	AACGTAGAGG	GCGGCACGT	AGGGTGTCAC TCCCACAGTG	

4751	GTTGCAAGAC (CAACGTTCTG (
4801	CGGAGGCCAT GCCTCCGGTA					·
4851	TTCGGCCCAT AAGCCGGGTA					
4901	TTTCATATGC AAAGTATACG					
4951	TGGACGACAC ACCTGCTGTG					
5001	CTTTGGGCCG GAAACCCGGC					
5051	CGGCTCCAAC GCCGAGGTTG				CGCCAGTAAC	
5101	ACTGGAGCGA TGACCTCGCT	GGCGATGTTC CCGCTACAAG	GGGGATTCCC CCCCTAAGGG	AATACGAGGT TTATGCTCCA	CGCCAACATC GCGGTTGTAG	
5151					CGCGCTACTT GCGCGATGAA	
5201					CGGGCGTATA GCCCGCATAT	
5251			GTTGAGATAG		ACTGCCGTTA	
5301 	AAGCTACTAC	GTCGAACCCG	CGTCCCAGCT	ACGCTGCGTT	TCGTCCGATC AGCAGGCTAG	
		TGACAGCCCG	CATGTGTTTA	GCGGGCGTCI	TCGCGCCGGC	
		ACCGACACAT	CTTCATGAGC	GGCTATCACC	TTTGGCTGCG	
		CAGGCTCCCG	TTTCCTTATC	TCATCTACGO	CTGGCCCTAG	
		TATTTTCTAF	A AATAAATCAG	AGGTCTTTT	CCCCCTTAC	
	TTTCTGGGGT	GGACATCCAA	A ACCGTTCGAT	CGAATTCAT	A CGCCATTTTG I GCGGTAAAAC	
	GTTCCGTACC	TTTTTATGTA	A TTGACTCTTA	A TCTCTTCAA		
5651	GTCCTTGTCT	ACCTTGTCG	A CTTATACCC	G GTTTGTCCT.	T ATCTGTGGTA A TAGACACCAT	

5701	AGCAGTTCCT TCGTCAAGGA	GCCCCGGCTC CGGGGCCGAG				
5751	TGGGCCAAAC ACCCGGTTTG	AGGATATCTG TCCTATAGAC				
5801	CAAGAACAGA GTTCTTGTCT	TGGTCCCCAG ACCAGGGGTC			CAAAGATCTC	
5851	AACCATCAGA TTGGTAGTCT	TGTTTCCAGG ACAAAGGTCC		CTGGACTTTA	CTGGGACACG	
5901 	CTTATTTGAA GAATAAACTT	CTAACCAATC GATTGGTTAG				
5951 		AGCTCAATAA TCGAGTTATT	TTCTCGGGTG	TTGGGGAGTG	AGCCCCGCGG	
6001		TTGACTGAGT AACTGACTCA	GCGGGCCCAT	GGGCACATAG	GTTATTTGGG	
60 51		GCATCCGACT CGTAGGCTGA				
61 01		TTGACTACCC AACTGATGGG				
2.1.3		TAATTTGGTT ATTAAACCAA	AAAAAAGAAT			
6201	ATAGTTGCAT TATCAACGTA	ATTACTTAGC	CGGTTGCGCG	CCCCTCTCCG	CCAAACGCAT	
6251		TCCGCTTCCT AGGCGAAGGA				
6301	GGCTGCGGCG CCGACGCCGC	AGCGGTATCA TCGCCATAGT				
6351	ACAGAATCAG TGTCTTAGTC	GGGATAACGC CCCTATTGCG				
6401		TTGGCATTTT	TCCGGCGCAA	CGACCGCAAA	AAGGTATCCG	
6451	AGGCGGGGG	TGACGAGCAT ACTGCTCGTA	GTGTTTTTAG	CTGCGAGTTC	AGTCTCCACC	
6501	GCTTTGGGCT	GTCCTGATAT	TTCTATGGTC	CGCAAAGGGG		
6551		AGAGGACAAG	GCTGGGACGG	CGAATGGCCT	ATGGACAGGC	,
6601		AAGCCCTTCG	CACCGCGAAA	GAGTATCGAG	TGCGACATCC	

6651		CGGTGTAGGT GCCACATCCA				
		CAGCCCGACC GTCGGGCTGG				
	AGTCCAACCC	GGTAAGACAC CCATTCTGTG				
6801		GCAGAGCGAG CGTCTCGCTC			TCAAGAACTT	
6851		AACTACGGCT TTGATGCCGA				
6901		GCCAGTTACC CGGTCAATGG				
6951 []	CCGTTTGTTT	CCACCGCTGG GGTGGCGACC	ATCGCCACCA		CGTTCGTCGT	
7001	CTAATGCGCG	AGAAAAAAAG TCTTTTTTC	CTAGAGTTCT	TCTAGGAAAC		
7051		CGCTCAGTGG GCGAGTCACC				
7101	ATGAGATTAT TACTCTAATA	CAAAAAGGAT GTTTTTCCTA	CTTCACCTAG GAAGTGGATC	ATCCTTTTGC TAGGAAAACG	GGCCGCAAAT CCGGCGTTTA	
7151	GTTAGATTTC		ATTTGAACCA		CCAATGCTTA GGTTACGAAT	
7201					CATCCATAGT GTAGGTATCA	
7251					G GGCTTACCAT C CCGAATGGTA	
7301					C ACCGGCTCCA G TGGCCGAGGT	
7351	CTAAATAGTC	GTTATTTGG1	CGGTCGGCCT	TCCCGGCTCC	C GCAGAAGTGG G CGTCTTCACC	
	AGGACGTTGA	A AATAGGCGGA	A GGTAGGTCAG	ATAATTAACA		
7451	GATCTCATT	TAGTTCGCCA	CAATTATCAA	ACGCGTTGC	I TGTTGCCATT A ACAACGGTAA	
	CGATGTCCG	T AGCACCACAC	G TGCGAGCAGC	C AAACCATAC		
7551	GAGGCCAAG	G GTTGCTAGT	T CCGCTCAATO	G TACTAGGGG	C ATGTTGTGCA G TACAACACGT	

7601	AAAAAGCGGT TTTTTCGCCA	TAGCTCCTTC ATCGAGGAAG	GGTCCTCCGA CCAGGAGGCT	TCGTTGTCAG AGCAACAGTC	AAGTAAGTTG TTCATTCAAC	
7651	GCCGCAGTGT CGGCGTCACA	TATCACTCAT ATAGTGAGTA	GGTTATGGCA CCAATACCGT	GCACTGCATA CGTGACGTAT	ATTCTCTTAC TAAGAGAATG	
7701	TGTCATGCCA ACAGTACGGT	TCCGTAAGAT AGGCATTCTA	GCTTTTCTGT CGAAAAGACA	GACTGGTGAG CTGACCACTC	TACTCAACCA ATGAGTTGGT	
7751	AGTCATTCTG TCAGTAAGAC	AGAATAGTGT TCTTATCACA	ATGCGGCGAC TACGCCGCTG	CGAGTTGCTC GCTCAACGAG	TTGCCCGGCG AACGGGCCGC	
7801	TCAATACGGG AGTTATGCCC	ATAATACCGC TATTATGGCG	GCCACATAGC CGGTGTATCG	AGAACTTTAA TCTTGAAATT	AAGTGCTCAT TTCACGAGTA	
7851	CATTGGAAAA GTAACCTTTT	CGTTCTTCGG GCAAGAAGCC	GGCGAAAACT CCGCTTTTGA	CTCAAGGATC GAGTTCCTAG	TTACCGCTGT AATGGCGACA	
79 01	TGAGATCCAG ACTCTAGGTC	TTCGATGTAA AAGCTACATT	CCCACTCGTG GGGTGAGCAC	CACCCAACTG GTGGGTTGAC	ATCTTCAGCA TAGAAGTCGT	
7951	TCTTTTACTT AGAAAATGAA	TCACCAGCGT AGTGGTCGCA	TTCTGGGTGA	GCAAAAACAG	GAAGGCAAAA CTTCCGTTTT	
8001	TGCCGCAAAA ACGGCGTTTT	AAGGGAATAA TTCCCTTATT	GGGCGACACG CCCGCTGTGC	GAAATGTTGA CTTTACAACT	ATACTCATAC TATGAGTATG	
8051	AGAAGGAAAA	TCAATATTAT AGTTATAATA	T TGAAGCATTI A ACTTCGTAAA	TAGTCCCAA	TTGTCTCATG AACAGAGTAC	
[8 101	AGCGGATACA TCGCCTATGT	ATAAACTTAG	G TATTTAGAAAC ATAAATCTT	A AATAAACAAA T TTATTTGTT	A TAGGGGTTCC C ATCCCCAAGG	
2 151	GCGCACATTT CGCGTGTAAA					

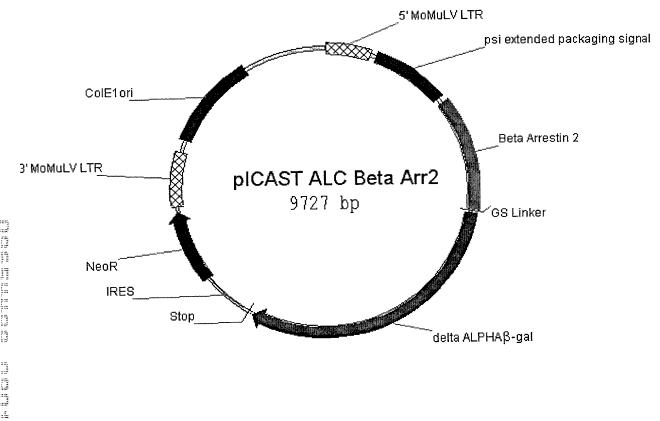


Figure 14

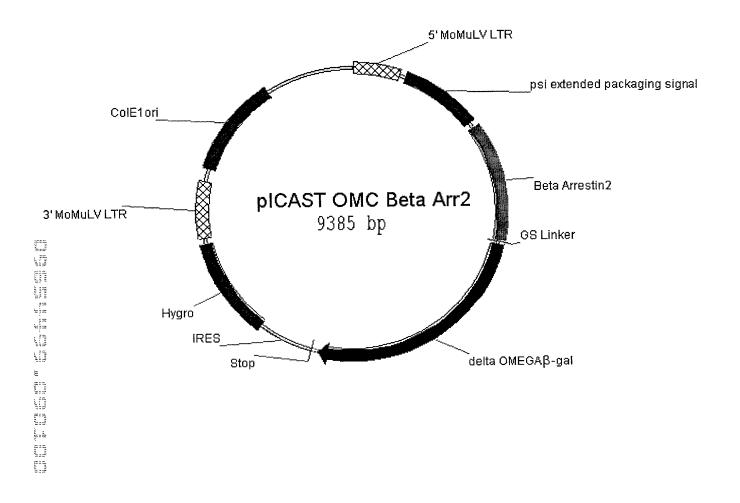


Figure 15

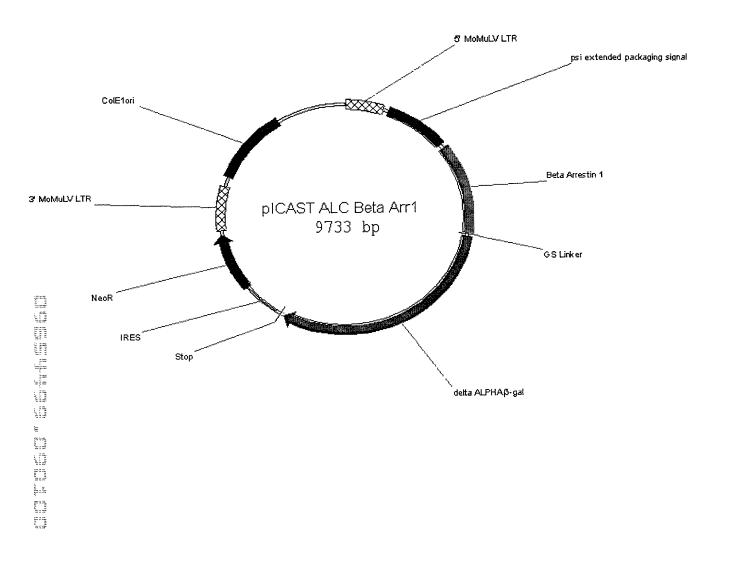


Figure 16

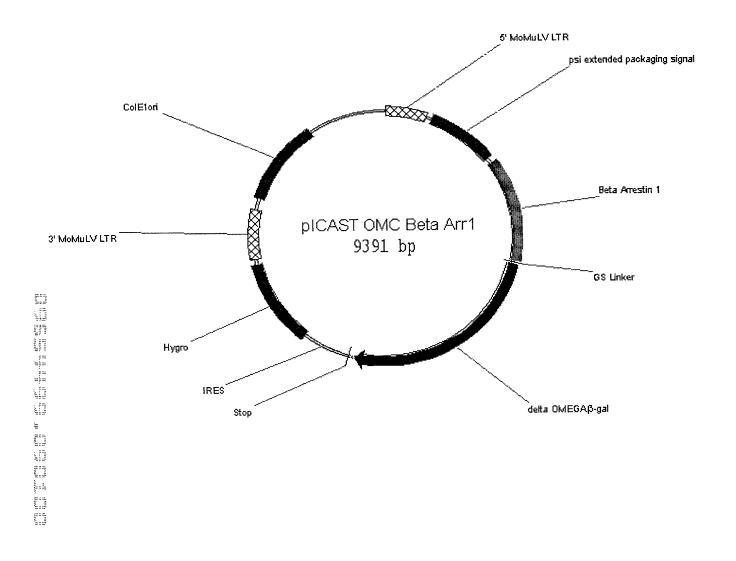


Figure 17

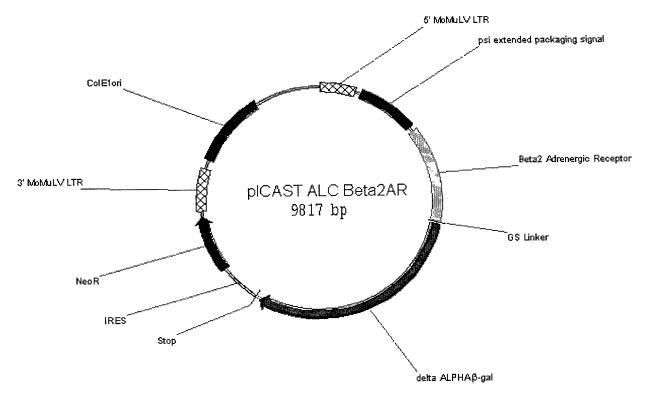


Figure 18

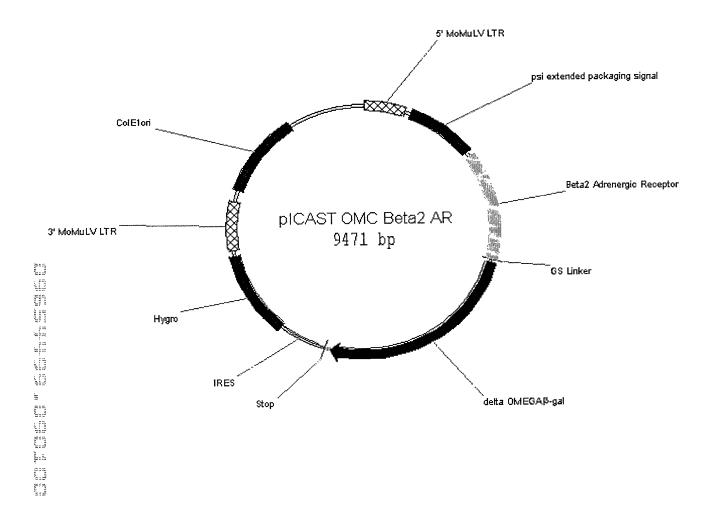


Figure 19

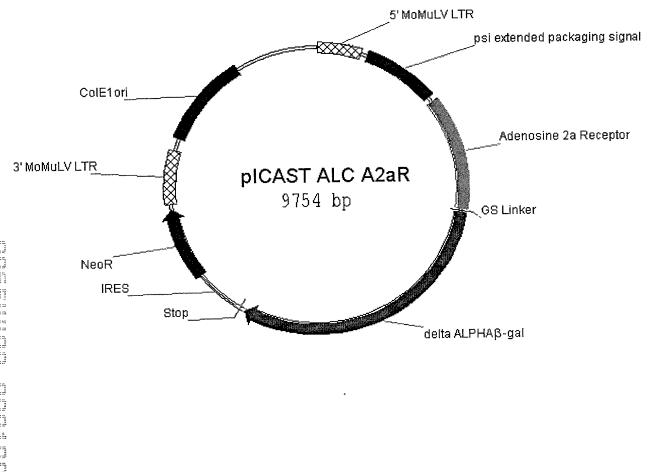


Figure 20

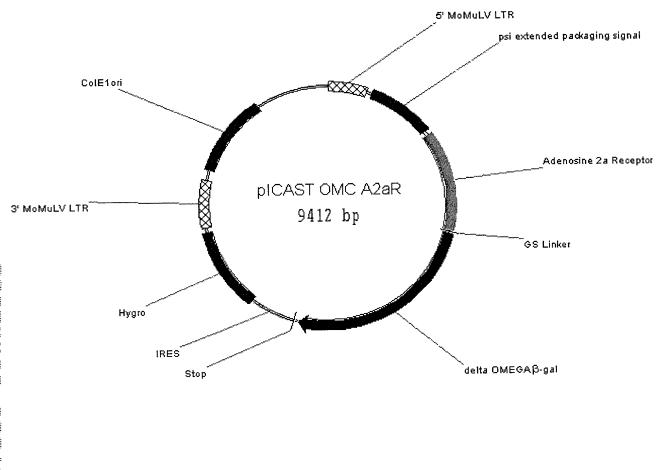


Figure 21

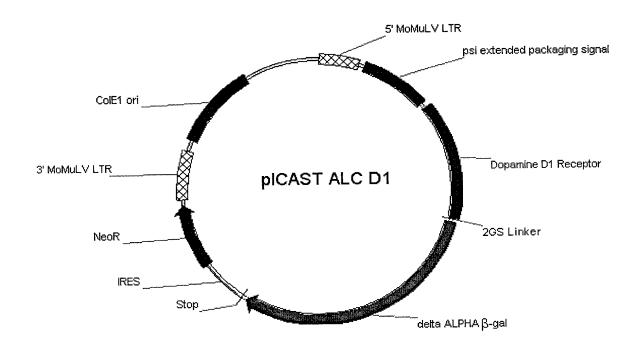


Figure 22

•

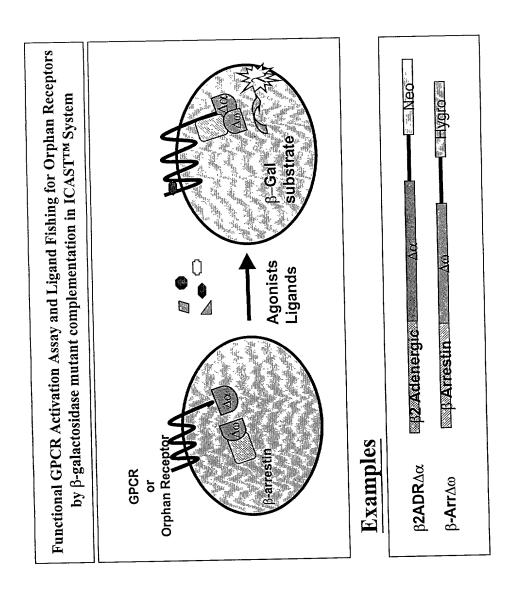


Figure 23

2.3

DOCKET NO. 4085-226-27

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Michelle A.J. PALMER, et al

ART UNIT:

SERIAL NO.: New Application

EXAMINER:

FILING DATE: Herewith

FOR:

RECEPTOR FUNCTION ASSAY FOR G-PROTEIN COUPLED

RECEPTORS AND ORPHAN RECEPTORS BY REPORTER ENZYME

MUTANT COMPLEMENTATION

LIST OF INVENTORS' NAMES AND ADDRESSES

ASSISTANT COMMISSIONER FOR PATENTS WASHINGTON, D.C. 20231

SIR:

Listed below are the names and addresses of the inventors for the above-identified patent application.

- 1) Michelle A.J. PALMER 87 Medford Street Arlington, MA 02174
- Melissa GEE3 North StreetGrafton, MA 01519
- 3) Bonnie TILLOTSON 4 Ripley Road Belmont, MA 02178
- 4) Xiao-Jia CHANG 25 Round Hill Road Lincoln, MA 01773

A declaration containing all the necessary information will be submitted at a later date.

Respectfully submitted,

PIPER MARBURY RUDNICK & WOLFE LLP

Steven B. Kelber

Registration No. 30,073

Attorney of Record

Laura A. Donnelly

Registration No. 38,435

1200 Nineteenth Street, N.W. Washington, D.C. 20036-2412 Telephone No. (202) 861-3900 Facsimile No. (202) 223-2085